GenCore version 5.1.6

OM nucleic - protein search, using frame_plus_n2p model

Run on: February 25, 2004, 05:40:34 ; Search time 87 Seconds
(without alignments)
11837.373 Million cell updates/sec

Title: US-09-938-956-6
Perfect score: 2908
Sequence: 1 atgagacgatccggaaacta..........tattcgagccctttgcatga 1632

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 2034082

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
```
-MODEL=frame+_n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US09938956/runat_24022004_144700_3071/app_query.fasta_1.1799
-DB=SPTREMBL_25 -QFMT=fastan -SUFFIX=rspt -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09938956_@CGN_1_1_116_@runat_24022004_144700_3071 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7
```

Database : SPTREMBL_25:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2876	98.9	599	10	Q40322	Q40322 mentha spic
2	2839	97.6	599	10	Q9SW76	Q9sw76 mentha long
3	2178	74.9	613	10	Q940E7	Q940e7 agastache r
4	2119	72.9	620	10	Q9FUW5	Q9fuw5 schizonepet
5	2007	69.0	604	10	Q9FV73	Q9fv73 perilla fru
6	2006	69.0	604	10	Q9FV75	Q9fv75 perilla cit
7	2006	69.0	604	10	Q9M5A4	Q9m5a4 perilla cit
8	2003.5	68.9	603	10	O04806	O04806 perilla fru
9	2002	68.8	604	10	Q9AXM7	Q9axm7 perilla fru
10	1998.5	68.7	603	10	Q9FV72	Q9fv72 perilla fru
11	1988.5	68.4	597	10	Q8L5J7	Q8l5j7 salvia sten
12	1937.5	66.6	608	10	Q9LKM5	Q9lkm5 perilla fru
13	1633.5	56.2	591	10	O81191	O81191 salvia offi
14	1553.5	53.4	590	10	O81193	O81193 salvia offi
15	1513.5	52.0	607	10	Q8W1J9	Q8w1j9 perilla fru
16	1468	50.5	598	10	O81192	O81192 salvia offi
17	1410.5	48.5	597	10	Q93X23	Q93x23 quercus ile
18	1332	45.8	606	10	Q8H2B4	Q8h2b4 mentha aqua
19	1207.5	41.5	583	10	Q9SPN1	Q9spn1 artemisia a
20	1202	41.3	606	10	Q8L5K1	Q8l5k1 citrus limo
21	1202	41.3	606	10	Q8L5K3	Q8l5k3 citrus limo
22	1200	41.3	600	10	Q8L5K4	Q8l5k4 citrus limo
23	1197.5	41.2	603	10	Q8GUE4	Q8gue4 cinnamomum
24	1196.5	41.1	602	10	Q8L5K2	Q8l5k2 citrus limo
25	1188.5	40.9	595	10	Q7XAS7	Q7xas7 populus tre
26	1179.5	40.6	567	10	Q9SPN0	Q9spn0 artemisia a
27	1176.5	40.5	595	10	Q9AR86	Q9ar86 populus x c
28	1168.5	40.2	582	10	Q94G53	Q94g53 artemisia a
29	1168.5	40.2	598	10	Q84UV1	Q84uv1 arabidopsis
30	1161.5	39.9	598	10	Q9LRZ6	Q9lrz6 arabidopsis
31	1157	39.8	600	10	Q9LDF1	Q9ldf1 arabidopsis
32	1147	39.4	555	10	Q9FVI8	Q9fvi8 arabidopsis
33	1147	39.4	591	10	Q9ZUH4	Q9zuh4 arabidopsis
34	1138.5	39.2	540	10	Q84NE5	Q84ne5 arabidopsis
35	1084.5	37.3	600	10	O23516	O23516 arabidopsis
36	1071	36.8	583	10	Q7Y1V1	Q7y1v1 melaleuca a
37	1039	35.7	1024	10	O23517	O23517 arabidopsis
38	995.5	34.2	576	10	Q84LB2	Q84lb2 malus domes
39	891.5	30.7	336	10	Q9FV74	Q9fv74 perilla fru
40	856	29.4	556	10	Q9FVL3	Q9fvl3 lycopersico
41	841	28.9	549	10	Q9XIZ0	Q9xiz0 solanum tub
42	840	28.9	551	10	Q9ZTQ6	Q9ztq6 solanum tub
43	837	28.8	549	10	Q9XJ25	Q9xj25 solanum tub
44	836	28.7	556	10	Q9XJ32	Q9xj32 solanum tub
45	835	28.7	560	10	Q9SDN9	Q9sdn9 capsicum an

ALIGNMENTS

RESULT 1
Q40322
```
ID   Q40322       PRELIMINARY;      PRT;   599 AA.
AC   Q40322;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   4S-limonene synthase precursor.
OS   Mentha spicata (spearmint).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC   lamiids; Lamiales; Lamiaceae; Nepetoideae; Mentheae; Mentha.
OX   NCBI_TaxID=29719;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Leaf;
RX   MEDLINE=94043077; PubMed=8226816;
RA   Colby S.M., Alonso W.R., Katahira E.J., McGarvey D.J., Croteau R.;
RT   "4S-limonene synthase from the oil glands of spearmint (Mentha
RT   spicata). cDNA isolation, characterization, and bacterial expression
RT   of the catalytically active monoterpene cyclase.";
RL   J. Biol. Chem. 268:23016-23024(1993).
DR   EMBL; L13459; AAC37366.1; -.
DR   PIR; A48863; A48863.
DR   HSSP; Q40577; 5EAU.
```

```
DR   GO; GO:0016829; F:lyase activity; IEA.
DR   GO; GO:0008152; P:metabolism; IEA.
DR   InterPro; IPR005630; Terpene_synth_C.
DR   InterPro; IPR008949; Terpenoid_synth.
DR   InterPro; IPR008930; Terp_cyc_toroid.
DR   InterPro; IPR001906; Terp_synth-like.
DR   Pfam; PF01397; Terpene_synth; 1.
DR   Pfam; PF03936; Terpene_synth_C; 1.
KW   Transit peptide.
FT   TRANSIT       1     99       POTENTIAL.
FT   CHAIN       100    599       4S-LIMONENE SYNTHASE.
SQ   SEQUENCE   599 AA;  69843 MW;  8BF53D23329CAF65 CRC64;

Alignment Scores:
Pred. No.:              6.64e-218    Length:         599
Score:                  2876.00      Matches:        542
Percent Similarity:     100.00%      Conservative:   0
Best Local Similarity:  100.00%      Mismatches:     0
Query Match:            98.90%       Indels:         0
DB:                     10           Gaps:           0

US-09-938-956-6 (1-1632) x Q40322 (1-599)

Qy          4 AGACGATCCGGAAACTACAACCCTTCTCGTTGGGATGTCAACTTCATCCAATCGCTTCTC 63
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         58 ArgArgSerGlyAsnTyrAsnProSerArgTrpAspValAsnPheIleGlnSerLeuLeu 77

Qy         64 AGTGACTATAAGGAGGACAAACACGTGATTAGGGCTTCTGAGCTGGTCACTTTGGTGAAG 123
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         78 SerAspTyrLysGluAspLysHisValIleArgAlaSerGluLeuValThrLeuValLys 97

Qy        124 ATGGAACTGGAGAAAGAAACGGATCAAATTCGACAACTTGAGTTGATCGATGACTTGCAG 183
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         98 MetGluLeuGluLysGluThrAspGlnIleArgGlnLeuGluLeuIleAspAspLeuGln 117

Qy        184 AGGATGGGGCTGTCCGATCATTTCCAAAATGAGTTCAAAGAAATCTTGTCCTCTATATAT 243
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        118 ArgMetGlyLeuSerAspHisPheGlnAsnGluPheLysGluIleLeuSerSerIleTyr 137

Qy        244 CTCGACCATCACTATTACAAGAACCCTTTTCCAAAAGAAGAAAGGGATCTCTACTCCACA 303
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        138 LeuAspHisHisTyrTyrLysAsnProPheProLysGluGluArgAspLeuTyrSerThr 157

Qy        304 TCTCTTGCATTTAGGCTCCTCAGAGAACATGGTTTTCAAGTCGCACAAGAGGTATTCGAT 363
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        158 SerLeuAlaPheArgLeuLeuArgGluHisGlyPheGlnValAlaGlnGluValPheAsp 177

Qy        364 AGTTTCAAGAACGAGGAGGGTGAGTTCAAAGAAAGCCTTAGCGACGACACCAGAGGATTG 423
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        178 SerPheLysAsnGluGluGlyGluPheLysGluSerLeuSerAspAspThrArgGlyLeu 197

Qy        424 TTGCAACTGTATGAAGCTTCCTTTCTGTTGACGGAAGGCGAAACCACGCTCGAGTCAGCG 483
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        198 LeuGlnLeuTyrGluAlaSerPheLeuLeuThrGluGlyGluThrThrLeuGluSerAla 217

Qy        484 AGGGAATTCGCCACCAAATTTTTGGAGGAAAAAGTGAACGAGGGTGGTGTTGATGGCGAC 543
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        218 ArgGluPheAlaThrLysPheLeuGluGluLysValAsnGluGlyGlyValAspGlyAsp 237

Qy        544 CTTTTAACAAGAATCGCATATTCTTTGGACATCCCTCTTCATTGGAGGATTAAAAGGCCA 603
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        238 LeuLeuThrArgIleAlaTyrSerLeuAspIleProLeuHisTrpArgIleLysArgPro 257

Qy        604 AATGCACCTGTGTGGATCGAATGGTATAGGAAGAGGCCCGACATGAATCCAGTAGTGTTG 663
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        258 AsnAlaProValTrpIleGluTrpTyrArgLysArgProAspMetAsnProValValLeu 277

Qy        664 GAGCTTGCCATACTCGACTTAAATATTGTTCAAGCACAATTTCAAGAAGAGCTCAAAGAA 723
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        278 GluLeuAlaIleLeuAspLeuAsnIleValGlnAlaGlnPheGlnGluGluLeuLysGlu 297

Qy        724 TCCTTCAGGTGGTGGAGAAATACTGGGTTTGTTGAGAAGCTGCCCTTCGCAAGGGATAGA 783
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        298 SerPheArgTrpTrpArgAsnThrGlyPheValGluLysLeuProPheAlaArgAspArg 317

Qy        784 CTGGTGGAATGCTACTTTTGGAATACTGGGATCATCGAGCCACGTCAGCATGCAAGTGCA 843
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        318 LeuValGluCysTyrPheTrpAsnThrGlyIleIleGluProArgGlnHisAlaSerAla 337

Qy        844 AGGATAATGATGGGCAAAGTCAACGCTCTGATTACGGTGATCGATGATATTTATGATGTC 903
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        338 ArgIleMetMetGlyLysValAsnAlaLeuIleThrValIleAspAspIleTyrAspVal 357

Qy        904 TATGGCACCTTAGAAGAACTCGAACAATTCACTGACCTCATTCGAAGATGGGATATAAAC 963
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        358 TyrGlyThrLeuGluGluLeuGluGlnPheThrAspLeuIleArgArgTrpAspIleAsn 377

Qy        964 TCAATCGACCAACTTCCCGATTACATGCAACTGTGCTTTCTTGCACTCAACAACTTCGTC 1023
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        378 SerIleAspGlnLeuProAspTyrMetGlnLeuCysPheLeuAlaLeuAsnAsnPheVal 397

Qy       1024 GATGATACATCGTACGATGTTATGAAGGAGAAAGGCGTCAACGTTATACCCTACCTGCGG 1083
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        398 AspAspThrSerTyrAspValMetLysGluLysGlyValAsnValIleProTyrLeuArg 417

Qy       1084 CAATCGTGGGTTGATTTGGCGGATAAGTATATGGTAGAGGCACGGTGGTTCTACGGCGGG 1143
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        418 GlnSerTrpValAspLeuAlaAspLysTyrMetValGluAlaArgTrpPheTyrGlyGly 437

Qy       1144 CACAAACCAAGTTTGGAAGAGTATTTGGAGAACTCATGGCAGTCGATAAGTGGGCCCTGT 1203
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        438 HisLysProSerLeuGluGluTyrLeuGluAsnSerTrpGlnSerIleSerGlyProCys 457

Qy       1204 ATGTTAACGCACATATTCTTCCGAGTAACAGATTCGTTCACAAAGGAGACCGTCGACAGT 1263
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        458 MetLeuThrHisIlePhePheArgValThrAspSerPheThrLysGluThrValAspSer 477

Qy       1264 TTGTACAAATACCACGATTTAGTTCGTTGGTCATCCTTCGTTCTGCGGCTTGCTGATGAT 1323
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        478 LeuTyrLysTyrHisAspLeuValArgTrpSerSerPheValLeuArgLeuAlaAspAsp 497

Qy       1324 TTGGGAACCTCGGTGGAAGAGGTGAGCAGAGGGGATGTGCCGAAATCACTTCAGTGCTAC 1383
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        498 LeuGlyThrSerValGluGluValSerArgGlyAspValProLysSerLeuGlnCysTyr 517

Qy       1384 ATGAGTGACTACAATGCATCGGAGGCGGAGGCGCGGAAGCACGTGAAATGGCTGATAGCG 1443
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        518 MetSerAspTyrAsnAlaSerGluAlaGluAlaArgLysHisValLysTrpLeuIleAla 537

Qy       1444 GAGGTGTGGAAGAAGATGAATGCGGAGAGGGTGTCGAAGGATTCTCCATTCGGCAAAGAT 1503
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        538 GluValTrpLysLysMetAsnAlaGluArgValSerLysAspSerProPheGlyLysAsp 557

Qy       1504 TTTATAGGATGTGCAGTTGATTTAGGAAGGATGGCGCAGTTGATGTACCATAATGGAGAT 1563
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        558 PheIleGlyCysAlaValAspLeuGlyArgMetAlaGlnLeuMetTyrHisAsnGlyAsp 577

Qy       1564 GGGCACGGCACACAACACCCTATTATACATCAACAAATGACCAGAACCTTATTCGAGCCC 1623
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        578 GlyHisGlyThrGlnHisProIleIleHisGlnGlnMetThrArgThrLeuPheGluPro 597

Qy       1624 TTTGCA 1629
              ||||||
Db        598 PheAla 599

RESULT 2
Q9SW76
ID   Q9SW76     PRELIMINARY;     PRT;    599 AA.
AC   Q9SW76;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Limonene synthase.
OS   Mentha longifolia (Horse mint).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
```

```
OC   lamiids; Lamiales; Lamiaceae; Nepetoideae; Mentheae; Mentha.
OX   NCBI_TaxID=38859;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Crock J., Croteau R.;
RT   "Limonene synthase cDNA from Mentha candicans.";
RL   Submitted (AUG-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF175323; AAD50304.1; -.
DR   HSSP; Q40577; 5EAU.
DR   GO; GO:0016829; F:lyase activity; IEA.
DR   GO; GO:0008152; P:metabolism; IEA.
DR   InterPro; IPR005630; Terpene_synth_C.
DR   InterPro; IPR008949; Terpenoid_synth.
DR   InterPro; IPR008930; Terp_cyc_toroid.
DR   InterPro; IPR001906; Terp_synth-like.
DR   Pfam; PF01397; Terpene_synth; 1.
DR   Pfam; PF03936; Terpene_synth_C; 1.
SQ   SEQUENCE   599 AA;  69905 MW;  BC789A0DD7D894EE CRC64;

Alignment Scores:
Pred. No.:              5.46e-215      Length:         599
Score:                  2839.00        Matches:        534
Percent Similarity:     98.89%         Conservative:   2
Best Local Similarity:  98.52%         Mismatches:     6
Query Match:            97.63%         Indels:         0
DB:                     10             Gaps:           0

US-09-938-956-6 (1-1632) x Q9SW76 (1-599)

Qy          4 AGACGATCCGGAAACTACAACCCTTCTCGTTGGGATGTCAACTTCATCCAATCGCTTCTC 63
              ||||||||||||||||||||||||||||||||||||||||||    |||||||||||||||
Db         58 ArgArgSerGlyAsnTyrAsnProSerArgTrpAspValGluPheIleGlnSerLeuHis 77

Qy         64 AGTGACTATAAGGAGGACAAACACGTGATTAGGGCTTCTGAGCTGGTCACTTTGGTGAAG 123
              ||||||||:::||||||||||||   ||||||||||||||||||||||||||||||||||||
Db         78 SerAspTyrGluGluAspLysHisAlaIleArgAlaSerGluLeuValThrLeuValLys 97

Qy        124 ATGGAACTGGAGAAAGAAACGGATCAAATTCGACAACTTGAGTTGATCGATGACTTGCAG 183
              ||||||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
Db         98 MetGluLeuGluLysGluThrAspHisIleArgGlnLeuGluLeuIleAspAspLeuGln 117

Qy        184 AGGATGGGGCTGTCCGATCATTTCCAAAATGAGTTCAAAGAAATCTTGTCCTCTATATAT 243
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        118 ArgMetGlyLeuSerAspHisPheGlnAsnGluPheLysGluIleLeuSerSerIleTyr 137

Qy        244 CTCGACCATCACTATTACAAGAACCCTTTTCCAAAAGAAGAAAGGGATCTCTACTCCACA 303
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        138 LeuAspHisHisTyrTyrLysAsnProPheProLysGluGluArgAspLeuTyrSerThr 157

Qy        304 TCTCTTGCATTTAGGCTCCTCAGAGAACATGGTTTTCAAGTCGCACAAGAGGTATTCGAT 363
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        158 SerLeuAlaPheArgLeuLeuArgGluHisGlyPheGlnValAlaGlnGluValPheAsp 177

Qy        364 AGTTTCAAGAACGAGGAGGGTGAGTTCAAAGAAAGCCTTAGCGACGACACCAGAGGATTG 423
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        178 SerPheLysAsnGluGluGlyGluPheLysGluSerLeuSerAspAspThrArgGlyLeu 197

Qy        424 TTGCAACTGTATGAAGCTTCCTTTCTGTTGACGGAAGGCGAAACCACGCTCGAGTCAGCG 483
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        198 LeuGlnLeuTyrGluAlaSerPheLeuLeuThrGluGlyGluThrThrLeuGluSerAla 217

Qy        484 AGGGAATTCGCCACCAAATTTTTGGAGGAAAAAGTGAACGAGGGTGGTGTTGATGGCGAC 543
              ||||||||||||||||||||||||||||||:::||||||||||||||||||||||||||||
Db        218 ArgGluPheAlaThrLysPheLeuGluGluArgValAsnGluGlyGlyValAspGlyAsp 237

Qy        544 CTTTTAACAAGAATCGCATATTCTTTGGACATCCCTCTTCATTGGAGGATTAAAAGGCCA 603
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        238 LeuLeuThrArgIleAlaTyrSerLeuAspIleProLeuHisTrpArgIleLysArgPro 257

Qy        604 AATGCACCTGTGTGGATCGAATGGTATAGGAAGAGGCCCGACATGAATCCAGTAGTGTTG 663
              ||||||||   ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        258 AsnAlaProAlaTrpIleGluTrpTyrArgLysArgProAspMetAsnProValValLeu 277

Qy        664 GAGCTTGCCATACTCGACTTAAATATTGTTCAAGCACAATTTCAAGAAGAGCTCAAAGAA 723
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        278 GluLeuAlaIleLeuAspLeuAsnIleValGlnAlaGlnPheGlnGluGluLeuLysGlu 297

Qy        724 TCCTTCAGGTGGTGGAGAAATACTGGGTTTGTTGAGAAGCTGCCCTTCGCAAGGGATAGA 783
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        298 SerPheArgTrpTrpArgAsnThrGlyPheValGluLysLeuProPheAlaArgAspArg 317

Qy        784 CTGGTGGAATGCTACTTTTGGAATACTGGGATCATCGAGCCACGTCAGCATGCAAGTGCA 843
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        318 LeuValGluCysTyrPheTrpAsnThrGlyIleIleGluProArgGlnHisAlaSerAla 337

Qy        844 AGGATAATGATGGGCAAAGTCAACGCTCTGATTACGGTGATCGATGATATTTATGATGTC 903
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        338 ArgIleMetMetGlyLysValAsnAlaLeuIleThrValIleAspAspIleTyrAspVal 357

Qy        904 TATGGCACCTTAGAAGAACTCGAACAATTCACTGACCTCATTCGAAGATGGGATATAAAC 963
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        358 TyrGlyThrLeuGluGluLeuGluGlnPheThrAspLeuIleArgArgTrpAspIleAsn 377

Qy        964 TCAATCGACCAACTTCCCGATTACATGCAACTGTGCTTTCTTGCACTCAACAACTTCGTC 1023
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        378 SerIleAspGlnLeuProAspTyrMetGlnLeuCysPheLeuAlaLeuAsnAsnPheVal 397

Qy       1024 GATGATACATCGTACGATGTTATGAAGGAGAAAGGCGTCAACGTTATACCCTACCTGCGG 1083
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        398 AspAspThrSerTyrAspValMetLysGluLysGlyValAsnValIleProTyrLeuArg 417

Qy       1084 CAATCGTGGGTTGATTTGGCGGATAAGTATATGGTAGAGGCACGGTGGTTCTACGGCGGG 1143
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        418 GlnSerTrpValAspLeuAlaAspLysTyrMetValGluAlaArgTrpPheTyrGlyGly 437

Qy       1144 CACAAACCAAGTTTGGAAGAGTATTTGGAGAACTCATGGCAGTCGATAAGTGGGCCCTGT 1203
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        438 HisLysProSerLeuGluGluTyrLeuGluAsnSerTrpGlnSerIleSerGlyProCys 457

Qy       1204 ATGTTAACGCACATATTCTTCCGAGTAACAGATTCGTTCACAAAGGAGACCGTCGACAGT 1263
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        458 MetLeuThrHisIlePhePheArgValThrAspSerPheThrLysGluThrValAspSer 477

Qy       1264 TTGTACAAATACCACGATTTAGTTCGTTGGTCATCCTTCGTTCTGCGGCTTGCTGATGAT 1323
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        478 LeuTyrLysTyrHisAspLeuValArgTrpSerSerPheValLeuArgLeuAlaAspAsp 497

Qy       1324 TTGGGAACCTCGGTGGAAGAGGTGAGCAGAGGGGATGTGCCGAAATCACTTCAGTGCTAC 1383
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        498 LeuGlyThrSerValGluGluValSerArgGlyAspValProLysSerLeuGlnCysTyr 517

Qy       1384 ATGAGTGACTACAATGCATCGGAGGCGGAGGCGCGGAAGCACGTGAAATGGCTGATAGCG 1443
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        518 MetSerAspTyrAsnAlaSerGluAlaGluAlaArgLysHisValLysTrpLeuIleAla 537

Qy       1444 GAGGTGTGGAAGAAGATGAATGCGGAGAGGGTGTCGAAGGATTCTCCATTCGGCAAAGAT 1503
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        538 GluValTrpLysLysMetAsnAlaGluArgValSerLysAspSerProPheGlyLysAsp 557

Qy       1504 TTTATAGGATGTGCAGTTGATTTAGGAAGGATGGCGCAGTTGATGTACCATAATGGAGAT 1563
              ||||||||||||||   ||||||||||||||||||||||||||||||||||||||||||||
Db        558 PheIleGlyCysAlaAlaAspLeuGlyArgMetAlaGlnLeuMetTyrHisAsnGlyAsp 577

Qy       1564 GGGCACGGCACACAACACCCTATTATACATCAACAAATGACCAGAACCTTATTCGAGCCC 1623
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        578 GlyHisGlyThrGlnHisProIleIleHisGlnGlnMetThrArgThrLeuPheGluPro 597

Qy       1624 TTTGCA 1629
              ||||||
Db        598 PheAla 599

RESULT 3
Q940E7
ID   Q940E7      PRELIMINARY;      PRT;   613 AA.
AC   Q940E7;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
```

```
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   D-limonene synthase.
OS   Agastache rugosa.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC   lamiids; Lamiales; Lamiaceae; Nepetoideae; Mentheae; Agastache.
OX   NCBI_TaxID=39271;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Saeki D., Maruyama T., Ito M., Honda G.;
RT   "Cloning and in vitro expression of d-limonene synthase from Agastache
RT   rugosa.";
RL   Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY055214; AAL17636.1; -.
DR   GO; GO:0016829; F:lyase activity; IEA.
DR   GO; GO:0008152; P:metabolism; IEA.
DR   InterPro; IPR005630; Terpene_synth_C.
DR   InterPro; IPR008949; Terpenoid_synth.
DR   InterPro; IPR008930; Terp_cyc_toroid.
DR   InterPro; IPR001906; Terp_synth-like.
DR   Pfam; PF01397; Terpene_synth; 1.
DR   Pfam; PF03936; Terpene_synth_C; 1.
SQ   SEQUENCE   613 AA;  71422 MW;  A92502C399F47572 CRC64;

Alignment Scores:
Pred. No.:               6.49e-163     Length:        613
Score:                   2178.00       Matches:       407
Percent Similarity:      86.40%        Conservative:  63
Best Local Similarity:   74.82%        Mismatches:    72
Query Match:             74.90%        Indels:        2
DB:                      10            Gaps:          2

US-09-938-956-6 (1-1632) x Q940E7 (1-613)

Qy          4 AGACGATCCGGAAACTACAACCCTTCTCGTTGGGATGTCAACTTCATCCAATCGCTTCTC 63
              ||||||||||||||||:::||||||||||||||||:::||||||||||||||
Db         70 ArgArgSerGlyAsnTyrSerProSerArgTrpAspValAspPheIleGlnSerLeuAsn 89

Qy         64 AGTGACTATAAGGAGGACAAACACGTGATTAGGGCTTCTGAGCTGGTCACTTTGGTGAAG 123
              ||||||||:::|||::::::|||      ||||||||||||||:::|||   ||||||
Db         90 SerAspTyrGlnGluGluArgHisThrArgArgAlaSerGluLeuIleThrGlnValLys 109

Qy        124 ATGGAACTGGAGAAAGAA---ACGGATCAAATTCGACAACTTGAGTTGATCGATGACTTG 180
              |||   :::||||||||||   ||||||    ||||||||||||||||||||||||||||
Db        110 MetLeuMetGluLysGluThrThrAspProIleArgGlnLeuGluLeuIleAspAspLeu 129

Qy        181 CAGAGGATGGGGCTGTCCGATCATTTCCAAAATGAGTTCAAAGAAATCTTGTCCTCTATA 240
              ||||||:::|||||||||||||||||||||||||||||||||||||||||:::::::|||
Db        130 GlnArgLeuGlyLeuSerAspHisPheGlnAsnGluPheLysGluIleLeuAsnThrIle 149

Qy        241 TATCTCGACCATCACTATTACAAG---AACCCTTTTCCAAAAGAAGAAAGGGATCTCTAC 297
              |||||||||:::    ||||||      |||         :::|||   |||||||||||
Db        150 TyrLeuAspAsnLysTyrTyrAsnIleAsnIleMetArgGluGluSerArgAspLeuTyr 169

Qy        298 TCCACATCTCTTGCATTTAGGCTCCTCAGAGAACATGGTTTTCAAGTCGCACAAGAGGTA 357
              ||||||::::||||||||||||||||||||||||||||||||||||||||||||||||||
Db        170 SerThrAlaLeuAlaPheArgLeuLeuArgGluHisGlyPheGlnValAlaGlnGluVal 189

Qy        358 TTCGATAGTTTCAAGAACGAGGAGGGTGAGTTCAAAGAAAGCCTTAGCGACGACACCAGA 417
              |||:::   ||||||||||||||||:::||||||   ||||||   ||||||||||||
Db        190 PheGluCysPheLysAsnGluGluGlyAspPheLysAlaSerLeuIleAspAspThrArg 209

Qy        418 GGATTGTTGCAACTGTATGAAGCTTCCTTTCTGTTGACGGAAGGCGAAACCACGCTCGAG 477
              |||||||||||||||||||||||||||||||||      |||||||||   |||||||||
Db        210 GlyLeuLeuGlnLeuTyrGluAlaSerPheLeuPheLysGluGlyGluAsnThrLeuGlu 229

Qy        478 TCAGCGAGGGAATTCGCCACCAAATTTTTGGAGGAAAAAGTGAACGAGGGTGGTGTTGAT 537
                    ||||||||||||    ||||||   |||:::||||||:::           :::|||
Db        230 IleAlaArgGluPheThrThrLysIleLeuGlnGluLysLeuLysGlyAspGluIleAsp 249

Qy        538 GGCGACCTTTTAACAAGAATCGCATATTCTTTGGACATCCCTCTTCATTGGAGGATTAAA 597
                 :::||||||:::   |||   ||||||||||:::||||||   :::|||   :::
Db        250 AspAsnLeuLeuSerSerIleArgTyrSerLeuGluIleProAsnTyrTrpSerValVal 269

Qy        598 AGGCCAAATGCACCTGTGTGGATCGAATGGTATAGGAAGAGGCCCGACATGAATCCAGTA 657
              |||||||||      |||||||||:::   ||||||||||||   ||||||||||||||||
Db        270 ArgProAsnValSerValTrpIleAspGluTyrArgLysArgSerAspMetAsnProVal 289

Qy        658 GTGTTGGAGCTTGCCATACTCGACTTAAATATTGTTCAAGCACAATTTCAAGAAGAGCTC 717
              ||||||||||||||||||||||||      ||||||||||||||||||   |||   ||||||
Db        290 ValLeuGluLeuAlaIleLeuAspAlaAsnIleValGlnAlaGlnLeuGlnLeuGluLeu 309

Qy        718 AAAGAATCCTTCAGGTGGTGGAGAAATACTGGGTTTGTTGAGAAGCTGCCCTTCGCAAGG 777
              |||||||||      |||||||||||||||||   ||||||||||||||||||||||||||||
Db        310 LysGluSerLeuArgTrpTrpArgAsnThrCysPheValGluLysLeuProPheAlaArg 329

Qy        778 GATAGACTGGTGGAATGCTACTTTTGGAATACTGGGATCATCGAGCCACGTCAGCATGCA 837
              |||||||||:::|||   |||||||||:::||||||::::::||||||||||||||||||
Db        330 AspArgLeuIleGluSerTyrPheTrpSerThrGlyMetValGluProArgGlnHisAla 349

Qy        838 AGTGCAAGGATAATGATGGGCAAAGTCAACGCTCTGATTACGGTGATCGATGATATTTAT 897
              :::|||||||||:::|||   ||||||   ||||||||||||||||:::||||||||||||
Db        350 AsnAlaArgIleIleMetAlaLysValIleAlaLeuIleThrValMetAspAspIleTyr 369

Qy        898 GATGTCTATGGCACCTTAGAAGAACTCGAACAATTCACTGACCTCATTCGAAGATGGGAT 957
              |||||||||||||||||||||||||||||||||||||:::          ||||||||||||
Db        370 AspValTyrGlyThrLeuGluGluLeuGluGlnPheThrGluAlaPheArgArgTrpAsp 389

Qy        958 ATAAACTCAATCGACCAACTTCCCGATTACATGCAACTGTGCTTTCTTGCACTCAACAAC 1017
              ::::::||||||||||||||||||   ||||||||||||||||||||||||:::||||||
Db        390 ValSerSerIleAspGlnLeuProThrTyrMetGlnLeuCysPheLeuAlaIleAsnAsn 409

Qy       1018 TTCGTCGATGATACATCGTACGATGTTATGAAGGAGAAAGGCGTCAACGTTATACCCTAC 1077
              ||||||||||||:::|||:::|||:::||||||   ||||||||||||:::      |||
Db        410 PheValAspAspThrAlaTyrAsnValLeuLysGluSerGlyValAsnValMetThrTyr 429

Qy       1078 CTGCGGCAATCGTGGGTTGATTTGGCGGATAAGTATATGGTAGAGGCACGGTGGTTCTAC 1137
              ||||||:::||||||||||||   |||:::   ||||||::::::|||::::::|||:::|||
Db        430 LeuArgLysSerTrpValAspGlnAlaGluAsnTyrLeuMetGluSerLysTrpTyrTyr 449

Qy       1138 GGCGGGCACAAACCAAGTTTGGAAGAGTATTTGGAGAACTCATGGCAGTCGATAAGTGGG 1197
                 ||||||||||||||||||||||||||||||||||||||      |||:::||||||
Db        450 SerGlyHisLysProSerLeuGluGluTyrLeuGluAsnSerTrpIleSerValSerGly 469

Qy       1198 CCCTGTATGTTAACGCACATATTCTTCCGAGTAACAGATTCGTTCACAAAGGAGACCGTC 1257
              ||||||:::|||||||||      ||||||   ||||||||||||      |||:::|||:::
Db        470 ProCysValLeuThrHisGluPhePheGlyValThrAspSerLeuAlaLysAspThrLeu 489

Qy       1258 GACAGTTTGTACAAATACCACGATTTAGTTCGTTGGTCATCCTTCGTTCTGCGGCTTGCT 1317
              ||||||||||||:::|||||||||||||:::|||||||::::::||||||||||||
Db        490 AspSerLeuTyrGluTyrHisAspIleValArgTrpSerSerTyrLeuLeuArgLeuAla 509

Qy       1318 GATGATTTGGGAACCTCGGTGGAAGAGGTGAGCAGAGGGGATGTGCCGAAATCACTTCAG 1377
              ||||||||||||||||||||||||||||||||||||||||||||||||||||:::|||
Db        510 AspAspLeuGlyThrSerValGluGluValSerArgGlyAspValProLysSerIleGln 529

Qy       1378 TGCTACATGAGTGACTACAATGCATCGGAGGCGGAGGCGCGGAAGCACGTGAAATGGCTG 1437
              |||||||||   |||   :::|||||||||   |||||||||:::|||:::|||   |||
Db        530 CysTyrMetHisAspAsnAspAlaSerGluGluGluAlaArgGlnHisIleLysGlyLeu 549

Qy       1438 ATAGCGGAGGTGTGGAAGAAGATGAATGCGGAGAGGGTGTCGAAGGATTCTCCATTCGGC 1497
              |||   |||:::||||||||||||||||||:::|||||||||||||
Db        550 IleArgGluMetTrpLysLysMetAsnValGluArgValSerGluAspSerProPheCys 569

Qy       1498 AAAGATTTTATAGGATGTGCAGTTGATTTAGGAAGGATGGCGCAGTTGATGTACCATAAT 1557
              :::|||||||||   |||      ||||||||||||||||||||   |||||||||
Db        570 ArgAspPheIleArgCysCysGluAspLeuGlyArgMetAlaGlnPheMetTyrHisTyr 589

Qy       1558 GGAGATGGGCACGGCACACAACACCCTATTATACATCAACAAATGACCAGAACCTTATTC 1617
              |||||||||||||||||||||||||   ||||||||||||:::         ||||||
Db        590 GlyAspGlyHisGlyThrGlnHisProLysIleHisGlnGlnIleAlaAlaCysLeuPhe 609

Qy       1618 GAGCCCTTTGCA 1629
              :::|||||||||
Db        610 GlnProPheAla 613
```

```
RESULT 4
Q9FUW5
ID   Q9FUW5      PRELIMINARY;      PRT;   620 AA.
AC   Q9FUW5;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   (+)-4R-limonene synthase.
OS   Schizonepeta tenuifolia.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC   lamiids; Lamiales; Lamiaceae; Nepetoideae; Mentheae; Schizonepeta.
OX   NCBI_TaxID=135200;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21200190; PubMed=11305598;
RA   Maruyama T., Ito M., Kiuchi F., Honda G.;
RT   "Molecular cloning, functional expression and characterization of d-
RT   limonene synthase from Schizonepeta tenuifolia.";
RL   Biol. Pharm. Bull. 24:373-377(2001).
DR   EMBL; AF282875; AAG01140.2; -.
DR   HSSP; Q40577; 5EAU.
DR   GO; GO:0016829; F:lyase activity; IEA.
DR   GO; GO:0008152; P:metabolism; IEA.
DR   InterPro; IPR005630; Terpene_synth_C.
DR   InterPro; IPR008949; Terpenoid_synth.
DR   InterPro; IPR008930; Terp_cyc_toroid.
DR   InterPro; IPR001906; Terp_synth-like.
DR   Pfam; PF01397; Terpene_synth; 1.
DR   Pfam; PF03936; Terpene_synth_C; 1.
SQ   SEQUENCE   620 AA;  71810 MW;  9AE8E7302442C55E CRC64;

Alignment Scores:
Pred. No.:               2.89e-158      Length:          620
Score:                   2119.00        Matches:         401
Percent Similarity:      84.31%         Conservative:    61
Best Local Similarity:   73.18%         Mismatches:      80
Query Match:             72.87%         Indels:          6
DB:                      10             Gaps:            3

US-09-938-956-6 (1-1632) x Q9FUW5 (1-620)

Qy          4 AGACGATCCGGAAACTACAACCCTTCTCGTTGGGATGTCAACTTCATCCAATCGCTTCTC 63

Db         73 ArgArgSerGlyAsnTyrLysProSerArgTrpAspValAspPheMetGlnSerLeuAsn 92

Qy         64 AGTGACTATAAGGAGGACAAACACGTGATTAGGGCTTCTGAGCTGGTCACTTTGGTGAAG 123

Db         93 SerAspTyrGlnGluGluArgHisArgThrLysAlaSerGluLeuIleThrGlnValLys 112

Qy        124 ATGGAACTGGAGAAAGAAACG------GATCAAATTCGACAACTTGAGTTGATCGATGAC 177

Db        113 AsnLeuLeuGluLysGluThrSerAspAspProIleArgGlnLeuGluLeuIleAspAsp 132

Qy        178 TTGCAGAGGATGGGGCTGTCCGATCATTTCCAAAATGAGTTCAAAGAAATCTTGTCCTCT 237

Db        133 LeuGlnArgLeuGlyLeuSerAspHisPheGluHisGluPheLysGluValLeuAsnSer 152

Qy        238 ATATATCTCGACCATCACTATTACAAGAACCCTTTTCCAAAAGAA---------GAAAGG 288

Db        153 IleTyrLeuAspAsnLysTyrTyrAsnIleAsnIleMetLysGluThrThrSerSerArg 172

Qy        289 GATCTCTACTCCACATCTCTTGCATTTAGGCTCCTCAGAGAACATGGTTTTCAAGTCGCA 348

Db        173 AspLeuTyrSerThrAlaLeuAlaPheArgLeuLeuArgGluHisGlyPheGlnValAla 192

Qy        349 CAAGAGGTATTCGATAGTTTCAAGAACGAGGAGGGTGAGTTCAAAGAAAGCCTTAGCGAC 408

Db        193 GlnGluValPheAspCysPheLysAsnGluGluGlyGluPheLysAlaSerLeuSerAsp 212

Qy        409 GACACCAGAGGATTGTTGCAACTGTATGAAGCTTCCTTTCTGTTGACGGAAGGCGAAACC 468

Db        213 AspProArgGlyLeuLeuGlnLeuTyrGluAlaSerPheLeuPheLysGluGlyGluAsn 232

Qy        469 ACGCTCGAGTCAGCGAGGGAATTCGCCACCAAATTTTTGGAGGAAAAAGTGAACGAGGGT 528

Db        233 ThrLeuGluIleAlaArgGluPheAlaThrLysLeuLeuGlnGluLysValAsnSerSer 252

Qy        529 GGT---GTTGATGGCGACCTTTTAACAAGAATCGCATATTCTTTGGACATCCCTCTTCAT 585

Db        253 AspGluIleAspAspAsnLeuLeuSerSerIleArgTyrSerLeuGluIleProThrTyr 272

Qy        586 TGGAGGATTAAAAGGCCAAATGCACCTGTGTGGATCGAATGGTATAGGAAGAGGCCCGAC 645

Db        273 TrpSerValIleArgProAsnValSerValTrpIleAspAlaTyrArgLysArgProAsp 292

Qy        646 ATGAATCCAGTAGTGTTGGAGCTTGCCATACTCGACTTAAATATTGTTCAAGCACAATTT 705

Db        293 MetAsnProValValLeuGluLeuAlaIleLeuAspAlaAsnIleMetGlnAlaGlnLeu 312

Qy        706 CAAGAAGAGCTCAAAGAATCCTTCAGGTGGTGGAGAAATACTGGGTTTGTTGAGAAGCTG 765

Db        313 GlnGlnGluLeuLysGluAlaLeuGlyTrpTrpArgAsnThrTrpPheValGluLysLeu 332

Qy        766 CCCTTCGCAAGGGATAGACTGGTGGAATGCTACTTTTGGAATACTGGGATCATCGAGCCA 825

Db        333 ProPheAlaArgAspArgLeuValGluSerTyrPheTrpSerThrGlyMetValProArg 352

Qy        826 CGTCAGCATGCAAGTGCAAGGATAATGATGGGCAAAGTCAACGCTCTGATTACGGTGATC 885

Db        353 ArgGlnHisLysThrAlaArgGlnLeuMetAlaLysValIleAlaLeuIleThrValMet 372

Qy        886 GATGATATTTATGATGTCTATGGCACCTTAGAAGAACTCGAACAATTCACTGACCTCATT 945

Db        373 AspAspIleTyrAspValTyrGlyThrLeuGluGluLeuGluLeuPheThrAspAlaPhe 392

Qy        946 CGAAGATGGGATATAAACTCAATCGACCAACTTCCCGATTACATGCAACTGTGCTTTCTT 1005

Db        393 ArgArgTrpAspValSerSerIleAspHisLeuProThrTyrMetGlnLeuCysPheLeu 412

Qy       1006 GCACTCAACAACTTCGTCGATGATACATCGTACGATGTTATGAAGGAGAAAGGCGTCAAC 1065

Db        413 SerIleAsnAsnPheValValAspThrAlaTyrAsnIleLeuLysGluThrGlyValAsn 432

Qy       1066 GTTATACCCTACCTGCGGCAATCGTGGGTTGATTTGGCGGATAAGTATATGGTAGAGGCA 1125

Db        433 ValThrThrTyrLeuGluLysSerTrpValAspGlnAlaGluAsnTyrLeuMetGluSer 452

Qy       1126 CGGTGGTTCTACGGCGGGCACAAACCAAGTTTGGAAGAGTATTTGGAGAACTCATGGCAG 1185

Db        453 LysTrpPheTyrSerGlyHisLysProSerLeuAspGluTyrLeuGluAsnSerTrpIle 472

Qy       1186 TCGATAAGTGGGCCCTGTATGTTAACGCACATATTCTTCCGAGTAACAGATTCGTTCACA 1245

Db        473 SerValSerGlyProCysValLeuThrHisGluPhePheGlyValThrAspSerLeuAla 492

Qy       1246 AAGGAGACCGTCGACAGTTTGTACAAATACCACGATTTAGTTCGTTGGTCATCCTTCGTT 1305

Db        493 LysAspThrLeuAspSerLeuTyrGluTyrHisAspIleValArgTrpSerSerTyrLeu 512

Qy       1306 CTGCGGCTTGCTGATGATTTGGGAACCTCGGTGGAAGAGGTGAGCAGAGGGGATGTGCCG 1365

Db        513 LeuArgLeuAlaAspAspLeuGlyThrSerValGluGluValSerArgGlyAspValPro 532

Qy       1366 AAATCACTTCAGTGCTACATGAGTGACTACAATGCATCGGAGGCGGAGGCGCGGAAGCAC 1425

Db        533 LysSerIleGlnCysTyrMetAsnAspAsnAsnAlaSerGluGluGluAlaArgGluHis 552

Qy       1426 GTGAAATGGCTGATAGCGGAGGTGTGGAAGAAGATGAATGCGGAGAGGGTGTCGAAGGAT 1485

Db        553 ValLysGlyLeuIleArgValMetTrpLysLysMetAsnAlaGluArgValSerGluAsp 572

Qy       1486 TCTCCATTCGGCAAAGATTTTATAGGATGTGCAGTTGATTTAGGAAGGATGGCGCAGTTG 1545

Db        573 SerProPheCysLysAspPheIleArgCysCysGluAspLeuGlyArgMetAlaGlnPhe 592
```

```
Qy       1546 ATGTACCATAATGGAGATGGGCACGGCACACAACACCCTATTATACATCAACAAATGACC 1605
              ||||||||  |||||||||||||||||||||||       |||||||||||:::|||
Db        593 MetTyrHisTyrGlyAspGlyHisGlyThrGlnHisAlaLysIleHisGlnGlnIleThr 612

Qy       1606 AGAACCTTATTCGAGCCCTTTGCA 1629
                      ||||||:::||||||||
Db        613 AspCysLeuPheGlnProPheAla 620

RESULT 5
Q9FV73
ID   Q9FV73      PRELIMINARY;      PRT;    604 AA.
AC   Q9FV73;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Limonene synthase.
OS   Perilla frutescens.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC   lamiids; Lamiales; Lamiaceae; Nepetoideae; Elsholtzieae; Perilla.
OX   NCBI_TaxID=48386;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=5598;
RA   Takeya H., Ito M., Dodo R., Honda G.;
RT   "Limonene synthase from Perilla sp.";
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF241792; AAG31437.1; -.
DR   HSSP; Q40577; 5EAT.
DR   GO; GO:0016829; F:lyase activity; IEA.
DR   GO; GO:0008152; P:metabolism; IEA.
DR   InterPro; IPR005630; Terpene_synth_C.
DR   InterPro; IPR008949; Terpenoid_synth.
DR   InterPro; IPR008930; Terp_cyc_toroid.
DR   InterPro; IPR001906; Terp_synth-like.
DR   Pfam; PF01397; Terpene_synth; 1.
DR   Pfam; PF03936; Terpene_synth_C; 1.
SQ   SEQUENCE   604 AA;  70569 MW;  1F1F9E3F7985815F CRC64;

Alignment Scores:
Pred. No.:                1.91e-149     Length:         604
Score:                    2007.00       Matches:        376
Percent Similarity:       82.05%        Conservative:   72
Best Local Similarity:    68.86%        Mismatches:     86
Query Match:              69.02%        Indels:         12
DB:                       10            Gaps:           3

US-09-938-956-6 (1-1632) x Q9FV73 (1-604)

Qy          4 AGACGATCCGGAAACTACAACCCTTCTCGTTGGGATGTCAACTTCATCCAATCGCTTCTC 63
              ||||||||||||||||||:::||||||   |||:::   ::::::|||    ||||||
Db         61 ArgArgSerGlyAsnTyrSerProSerPheTrpAsnAlaAspTyrIleLeuSerLeuAsn 80

Qy         64 AGTGACTATAAGGAGGAC---AAACACGTGATTAGGGCTTCTGAGCTGGTCACTTTGGTG 120
              :::   |||||||||:::   :::|||:::   ||||||  ||||||:::        |||
Db         81 AsnHisTyrLysGluGluSerArgHisMetLysArgAlaGlyGluLeuIleValGlnVal 100

Qy        121 AAGATGGAACTGGAGAAAGAAACGGATCAAATTCGACAACTTGAGTTGATCGATGACTTG 180
              ||||||    :::   |||||||||||    :::   |||||||||||||||||||||||
Db        101 LysMetValMetGlyLysGluThrAspProValValGlnLeuGluLeuIleAspAspLeu 120

Qy        181 CAGAGGATGGGGCTGTCCGATCATTTCCAAAATGAGTTCAAAGAAATCTTGTCCTCTATA 240
                 ::::::  ||||||  ||||||:::  |||   |||||||||||     :::|||
Db        121 HisLysLeuAlaLeuSerHisHisPheGluLysGluIleLysGluIleLeuPheAsnIle 140

Qy        241 TATCTCGACCATCACTATTACAAGAACCCTTTTCCAAAAGAAGAAAGGGATCTCTACTCC 300
                 :::       |||                         ||||||||||||||||||
Db        141 SerIleTyrAspHisLysIle---------------MetValGluArgAspLeuTyrSer 155

Qy        301 ACATCTCTTGCATTTAGGCTCCTCAGAGAACATGGTTTTCAAGTCGCACAAGAGGTATTC 360
              |||:::||||||||||||||||||||||:::||||||:::||||||   ||||||||||||
Db        156 ThrAlaLeuAlaPheArgLeuLeuArgGlnHisGlyTyrGlnValHisGlnGluValPhe 175

Qy        361 GATAGTTTCAAGAACGAGGAGGGTGAGTTCAAAGAAAGCCTTAGCGACGACACCAGAGGA 420
              |||   |||||||||:::   |||||||||||    |||||||||     ||||||:::|||
Db        176 AspCysPheLysAsnAspAsnGlyGluPheLysArgSerLeuSerSerAspThrLysGly 195

Qy        421 TTGTTGCAACTGTATGAAGCTTCCTTTCTGTTGACGGAAGGCGAAACCACGCTCGAGTCA 480
              |||||||||||||||||||||||||||||||||||||||||||||||   ||||||||||
Db        196 LeuLeuGlnLeuTyrGluAlaSerPheLeuLeuThrGluGlyGluMetThrLeuGluLeu 215

Qy        481 GCGAGGGAATTCGCCACCAAATTTTTGGAGGAAAAAGTGAACGAGGGTGGTGTTGATGGC 540
              ||||||||||||||||||   ||||||:::||||||:::|||:::       :::|||
Db        216 AlaArgGluPheAlaThrIlePheLeuGlnGluLysLeuAsnAspLysThrIleAspAsp 235

Qy        541 GAC-----------------CTTTTAACAAGAATCGCATATTCTTTGGACATCCCTCTT 582
              |||                 |||::::::   :::   :::||||||||||||||||:::
Db        236 AspAspAspAlaAspThrAsnLeuIleSerCysValArgHisSerLeuAspIleProIle 255

Qy        583 CATTGGAGGATTAAAAGGCCAAATGCACCTGTGTGGATCGAATGGTATAGGAAGAGGCCC 642
              ||||||||||||:::||||||||||||         ||||||:::    |||:::::|||
Db        256 HisTrpArgIleGlnArgProAsnAlaSerTrpTrpIleAspAlaTyrLysArgArgSer 275

Qy        643 GACATGAATCCAGTAGTGTTGGAGCTTGCCATACTCGACTTAAATATTGTTCAAGCACAA 702
                 |||||||||:::||||||||||||||||   ||||||||||||||   ||||||||
Db        276 HisMetAsnProLeuValLeuGluLeuAlaLysLeuAspLeuAsnIlePheGlnAlaGln 295

Qy        703 TTTCAAGAAGAGCTCAAAGAATCCTTCAGGTGGTGGAGAAATACTGGGTTTGTTGAGAAG 762
              ||||||:::|||||||||:::               ||||||:::||||||         ||||||
Db        296 PheGlnGlnGluLeuLysGlnAspLeuGlyTrpTrpLysAsnThrCysLeuAlaGluLys 315

Qy        763 CTGCCCTTCGCAAGGGATAGACTGGTGGAATGCTACTTTTGGAATACTGGGATCATCGAG 822
              ||||||||||||||||||||||||||||||||||||||||||   ||||||||||||:::
Db        316 LeuProPheAlaArgAspArgLeuValGluCysTyrPheTrpCysThrGlyIleIleGln 335

Qy        823 CCACGTCAGCATGCAAGTGCAAGGATAATGATGGGCAAAGTCAACGCTCTGATTACGGTG 882
              |||   ||||||   :::||||||:::   :::   ||||||||||||||||||||||
Db        336 ProLeuGlnHisGluAsnAlaArgValThrLeuAlaLysValAsnAlaLeuIleThrThr 355

Qy        883 ATCGATGATATTTATGATGTCTATGGCACCTTAGAAGAACTCGAACAATTCACTGACCTC 942
              :::||||||||||||||||||||||||||||||||||||||||||   ||||||:::
Db        356 LeuAspAspIleTyrAspValTyrGlyThrLeuGluGluLeuGluLeuPheThrGluAla 375

Qy        943 ATTCGAAGATGGGATATAAACTCAATCGACCAACTTCCCGATTACATGCAACTGTGCTTT 1002
              |||||||||||||||:::   |||||||||   |||||||:::||||||||||||||||||
Db        376 IleArgArgTrpAspValIleSerIleAspHisLeuProAsnTyrMetGlnLeuCysPhe 395

Qy       1003 CTTGCACTCAACAACTTCGTCGATGATACATCGTACGATGTTATGAAGGAGAAAGGCGTC 1062
              |||||||||||||||||||||||||||:::||||||||||||||||||||||||   :::
Db        396 LeuAlaLeuAsnAsnPheValAspAspThrAlaTyrAspValMetLysGluLysAspIle 415

Qy       1063 AACGTTATACCCTACCTGCGGCAATCGTGGGTTGATTTGGCGGATAAGTATATGGTAGAG 1122
              |||:::||||||||||||||||:::||||||:::||||||||||||:::   |||:::|||
Db        416 AsnIleIleProTyrLeuArgLysSerTrpLeuAspLeuAlaGluThrTyrLeuValGlu 435

Qy       1123 GCACGGTGGTTCTACGGCGGGCACAAACCAAGTTTGGAAGAGTATTTGGAGAACTCATGG 1182
              |||:::|||||||||   ||||||||||||::::::|||||||||||||   |||:::|||
Db        436 AlaLysTrpPheTyrSerGlyHisLysProAsnMetGluGluTyrLeuAsnAsnAlaTrp 455

Qy       1183 CAGTCGATAAGTGGGCCCTGTATGTTAACGCACATATTCTTCCGAGTAACAGATTCGTTC 1242
                 ||||||||||||||   ||||||   |||:::||||||||||||||||||||||
Db        456 IleSerIleSerGlyProValMetLeuCysHisValPhePheArgValThrAspSerIle 475

Qy       1243 ACAAAGGAGACCGTCGACAGTTTGTACAAATACCACGATTTAGTTCGTTGGTCATCCTTC 1302
              |||:::|||||||||:::||||||:::||||||||||||||||||:::|||:::|||||
Db        476 ThrArgGluThrValGluSerLeuPheLysTyrHisAspLeuIleArgTyrSerSerThr 495

Qy       1303 GTTCTGCGGCTTGCTGATGATTTGGGAACCTCGGTGGAAGAGGTGAGCAGAGGGGATGTG 1362
              :::||||||||||||||||||||||||:::|||||||||||||||||||||||||||
Db        496 IleLeuArgLeuAlaAspAspLeuGlyThrSerLeuGluGluValSerArgGlyAspVal 515

Qy       1363 CCGAAATCACTTCAGTGCTACATGAGTGACTACAATGCATCGGAGGCGGAGGCGCGGAAG 1422
              |||||||||:::||||||||||||:::|||   ||||||||||||   ||||||||:::
Db        516 ProLysSerIleGlnCysTyrMetAsnAspAspAsnAlaSerGluGluGluAlaArgArg 535
```

```
Qy     1423 CACGTGAAATGGCTGATAGCGGAGGTGTGGAAGAAGATGAATGCGGAGAGGGTGTCGAAG 1482
Db      536 HisValArgTrpLeuIleAlaGluThrTrpLysLysIleAsnGluGluValTrpSerAla 555

Qy     1483 GATTCTCCATTCGGCAAAGATTTTATAGGATGTGCAGTTGATTTAGGAAGGATGGCGCAG 1542
Db      556 AspSerProPheCysLysAspPheIleAlaCysAlaAlaAspMetGlyArgMetAlaGln 575

Qy     1543 TTGATGTACCATAATGGAGATGGGCACGGCACACAACACCCTATTATACATCAACAAATG 1602
Db      576 PheMetTyrHisAsnGlyAspGlyHisGlyIleGlnAsnProGlnIleHisGlnGlnMet 595

Qy     1603 ACCAGAACCTTATTCGAG 1620
Db      596 ThrAspIleLeuPheGlu 601
```

```
RESULT 6
Q9FV75
ID   Q9FV75      PRELIMINARY;      PRT;    604 AA.
AC   Q9FV75;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Limonene synthase.
OS   Perilla citriodora.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC   lamiids; Lamiales; Lamiaceae; Nepetoideae; Elsholtzieae; Perilla.
OX   NCBI_TaxID=121094;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=5601;
RA   Dodo R., Ito M., Takeya H., Honda G.;
RT   "Limonene synthase from Perilla sp.";
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF241790; AAG31435.1; -.
DR   HSSP; Q40577; 5EAT.
DR   GO; GO:0016829; F:lyase activity; IEA.
DR   GO; GO:0008152; P:metabolism; IEA.
DR   InterPro; IPR005630; Terpene_synth_C.
DR   InterPro; IPR008949; Terpenoid_synth.
DR   InterPro; IPR008930; Terp_cyc_toroid.
DR   InterPro; IPR001906; Terp_synth-like.
DR   Pfam; PF01397; Terpene_synth; 1.
DR   Pfam; PF03936; Terpene_synth_C; 1.
SQ   SEQUENCE   604 AA;  70498 MW;  60363DE340721A1C CRC64;

Alignment Scores:
Pred. No.:              2.29e-149     Length:          604
Score:                  2006.00       Matches:         375
Percent Similarity:     82.23%        Conservative:    74
Best Local Similarity:  68.68%        Mismatches:      85
Query Match:            68.98%        Indels:          12
DB:                     10            Gaps:            3

US-09-938-956-6 (1-1632) x Q9FV75 (1-604)
```

```
Qy        4 AGACGATCCGGAAACTACAACCCTTCTCGTTGGGATGTCAACTTCATCCAATCGCTTCTC 63
Db       61 ArgArgSerGlyAsnTyrSerProSerPheTrpAsnAlaAspTyrIleLeuSerLeuAsn 80

Qy       64 AGTGACTATAAGGAGGAC---AAACACGTGATTAGGGCTTCTGAGCTGGTCACTTTGGTG 120
Db       81 AsnHisTyrLysGluGluSerArgHisMetLysArgAlaGlyGluLeuIleValGlnVal 100

Qy      121 AAGATGGAACTGGAGAAAGAAACGGATCAAATTCGACAACTTGAGTTGATCGATGACTTG 180
Db      101 LysMetValMetGlyLysGluThrAspProValValGlnLeuGluLeuIleAspAspLeu 120

Qy      181 CAGAGGATGGGGCTGTCCGATCATTTCCAAAATGAGTTCAAAGAAATCTTGTCCTCTATA 240
Db      121 HisLysLeuAlaLeuSerHisHisPheGluLysGluIleLysGluIleLeuPheAsnIle 140

Qy      241 TATCTCGACCATCACTATTACAAGAACCCTTTTCCAAAAGAAGAAAGGGATCTCTACTCC 300
Db      141 SerIleTyrAspHisLysIle---------------MetValGluArgAspLeuTyrSer 155

Qy      301 ACATCTCTTGCATTTAGGCTCCTCAGAGAACATGGTTTTCAAGTCGCACAAGAGGTATTC 360
Db      156 ThrAlaLeuAlaPheArgLeuLeuArgGlnTyrGlyPheLysValProGlnGluValPhe 175

Qy      361 GATAGTTTCAAGAACGAGGAGGGTGAGTTCAAAGAAAGCCTTAGCGACGACACCAGAGGA 420
Db      176 AspCysPheLysAsnAspAsnGlyGluPheLysArgSerLeuSerSerAspThrLysGly 195

Qy      421 TTGTTGCAACTGTATGAAGCTTCCTTTCTGTTGACGGAAGGCGAAACCACGCTCGAGTCA 480
Db      196 LeuLeuGlnLeuTyrGluAlaSerPheLeuLeuThrGluGlyGluMetThrLeuGluLeu 215

Qy      481 GCGAGGGAATTCGCCACCAAATTTTTGGAGGAAAAAGTGAACGAGGGTGGTGTTGATGGC 540
Db      216 AlaArgGluPheAlaThrIlePheLeuGlnGluLysLeuAsnAspLysThrIleAspAsp 235

Qy      541 GAC------------------CTTTTAACAAGAATCGCATATTCTTTGGACATCCCTCTT 582
Db      236 AspAspAspAlaAspThrAsnLeuIleSerCysValArgHisSerLeuAspIleProIle 255

Qy      583 CATTGGAGGATTAAAAGGCCAAATGCACCTGTGTGGATCGAATGGTATAGGAAGAGGCCC 642
Db      256 HisTrpArgIleGlnArgProAsnAlaSerTrpTrpIleAspAlaTyrLysArgArgSer 275

Qy      643 GACATGAATCCAGTAGTGTTGGAGCTTGCCATACTCGACTTAAATATTGTTCAAGCACAA 702
Db      276 HisMetAsnProLeuValLeuGluLeuAlaLysLeuAspLeuAsnIlePheGlnAlaGln 295

Qy      703 TTTCAAGAAGAGCTCAAAGAATCCTTCAGGTGGTGGAGAAATACTGGGTTTGTTGAGAAG 762
Db      296 PheGlnGlnGluLeuLysGlnAspLeuGlyTrpTrpLysAsnThrCysLeuAlaGluLys 315

Qy      763 CTGCCCTTCGCAAGGGATAGACTGGTGGAATGCTACTTTTGGAATACTGGGATCATCGAG 822
Db      316 LeuProPheAlaArgAspArgLeuValGluCysTyrPheTrpCysThrGlyIleIleGln 335

Qy      823 CCACGTCAGCATGCAAGTGCAAGGATAATGATGGGCAAAGTCAACGCTCTGATTACGGTG 882
Db      336 ProLeuGlnHisGluAsnAlaArgValThrLeuAlaLysValAsnAlaLeuIleThrThr 355

Qy      883 ATCGATGATATTTATGATGTCTATGGCACCTTAGAAGAACTCGAACAATTCACTGACCTC 942
Db      356 LeuAspAspIleTyrAspValTyrGlyThrLeuGluGluLeuGluLeuPheThrGluAla 375

Qy      943 ATTCGAAGATGGGATATAAACTCAATCGACCAACTTCCCGATTACATGCAACTGTGCTTT 1002
Db      376 IleArgArgTrpAspValSerSerIleAspHisLeuProAsnTyrMetGlnLeuCysPhe 395

Qy     1003 CTTGCACTCAACAACTTCGTCGATGATACATCGTACGATGTTATGAAGGAGAAAGGCGTC 1062
Db      396 LeuAlaLeuAsnAsnPheValAspAspThrAlaTyrAspValMetLysGluLysAspIle 415

Qy     1063 AACGTTATACCCTACCTGCGGCAATCGTGGGTTGATTTGGCGGATAAGTATATGGTAGAG 1122
Db      416 AsnIleIleProTyrLeuArgLysSerTrpLeuAspLeuAlaGluThrTyrLeuValGlu 435

Qy     1123 GCACGGTGGTTCTACGGCGGGCACAAACCAAGTTTGGAAGAGTATTTGGAGAACTCATGG 1182
Db      436 AlaLysTrpPheTyrSerGlyHisLysProAsnMetGluGluTyrLeuAsnAsnAlaTrp 455

Qy     1183 CAGTCGATAAGTGGGCCCTGTATGTTAACGCACATATTCTTCCGAGTAACAGATTCGTTC 1242
Db      456 IleSerIleSerGlyProValMetLeuCysHisValPhePheArgValThrAspSerIle 475

Qy     1243 ACAAAGGAGACCGTCGACAGTTTGTACAAATACCACGATTTAGTTCGTTGGTCATCCTTC 1302
Db      476 ThrArgGluThrValGluSerLeuPheLysTyrHisAspLeuIleArgTyrSerSerThr 495

Qy     1303 GTTCTGCGGCTTGCTGATGATTTGGGAACCTCGGTGGAAGAGGTGAGCAGAGGGGATGTG 1362
```

```
Db     496 IleLeuArgLeuAlaAspAspLeuGlyThrSerLeuGluGluValSerArgGlyAspVal 515

Qy    1363 CCGAAATCACTTCAGTGCTACATGAGTGACTACAATGCATCGGAGGCGGAGGCGCGGAAG 1422
Db     516 ProLysSerIleGlnCysTyrMetAsnAspAsnAsnAlaSerGluGluGluAlaArgArg 535

Qy    1423 CACGTGAAATGGCTGATAGCGGAGGTGTGGAAGAAGATGAATGCGGAGAGGGTGTCGAAG 1482
Db     536 HisValArgTrpLeuIleAlaGluThrTrpLysLysIleAsnGluGluValTrpSerAla 555

Qy    1483 GATTCTCCATTCGGCAAAGATTTTATAGGATGTGCAGTTGATTTAGGAAGGATGGCGCAG 1542
Db     556 AspSerProPheCysLysAspPheIleAlaCysAlaAlaAspMetGlyArgMetAlaGln 575

Qy    1543 TTGATGTACCATAATGGAGATGGGCACGGCACACAACACCCTATTATACATCAACAAATG 1602
Db     576 PheMetTyrHisAsnGlyAspGlyHisGlyIleGlnAsnProGlnIleHisGlnGlnMet 595

Qy    1603 ACCAGAACCTTATTCGAG 1620
Db     596 ThrAspIleLeuPheGlu 601

RESULT 7
Q9M5A4
ID   Q9M5A4      PRELIMINARY;      PRT;   604 AA.
AC   Q9M5A4;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Limonene synthase.
OS   Perilla citriodora.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC   lamiids; Lamiales; Lamiaceae; Nepetoideae; Elsholtzieae; Perilla.
OX   NCBI_TaxID=121094;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20189430; PubMed=10726896;
RA   Ito M., Kiuchi F., Yang L.L., Honda G.;
RT   "Perilla citriodora from Taiwan and its phytochemical
RT   characteristics.";
RL   Biol. Pharm. Bull. 23:359-362(2000).
DR   EMBL; AF233894; AAF65545.1; -.
DR   HSSP; Q40577; 5EAT.
DR   GO; GO:0016829; F:lyase activity; IEA.
DR   GO; GO:0008152; P:metabolism; IEA.
DR   InterPro; IPR005630; Terpene_synth_C.
DR   InterPro; IPR008949; Terpenoid_synth.
DR   InterPro; IPR008930; Terp_cyc_toroid.
DR   InterPro; IPR001906; Terp_synth-like.
DR   Pfam; PF01397; Terpene_synth; 1.
DR   Pfam; PF03936; Terpene_synth_C; 1.
SQ   SEQUENCE   604 AA;  70597 MW;  6A3046E7D3A4C129 CRC64;

Alignment Scores:
Pred. No.:               2.29e-149      Length:         604
Score:                   2006.00        Matches:        375
Percent Similarity:      82.23%         Conservative:   74
Best Local Similarity:   68.68%         Mismatches:     85
Query Match:             68.98%         Indels:         12
DB:                      10             Gaps:           3

US-09-938-956-6 (1-1632) x Q9M5A4 (1-604)

Qy       4 AGACGATCCGGAAACTACAACCCTTCTCGTTGGGATGTCAACTTCATCCAATCGCTTCTC 63
Db      61 ArgArgSerGlyAsnTyrSerProSerPheTrpAsnAlaAspTyrIleLeuSerLeuAsn 80

Qy      64 AGTGACTATAAGGAGGAC---AAACACGTGATTAGGGCTTCTGAGCTGGTCACTTTGGTG 120
Db      81 AsnHisTyrLysGluGluSerArgHisMetLysArgAlaGlyGluLeuIleValGlnVal 100

Qy     121 AAGATGGAACTGGAGAAAGAAACGGATCAAATTCGACAACTTGAGTTGATCGATGACTTG 180
Db     101 LysMetValMetGlyLysGluThrAspProValValGlnLeuGluLeuIleAspAspLeu 120

Qy     181 CAGAGGATGGGGCTGTCCGATCATTTCCAAAATGAGTTCAAAGAAATCTTGTCCTCTATA 240
Db     121 HisLysLeuAlaLeuSerHisHisPheGluLysGluIleLysGluIleLeuPheAsnIle 140

Qy     241 TATCTCGACCATCACTATTACAAGAACCCTTTTCCAAAAGAAGAAAGGGATCTCTACTCC 300
Db     141 SerIleTyrAspHisLysIle---------------MetValGluArgAspLeuTyrSer 155

Qy     301 ACATCTCTTGCATTTAGGCTCCTCAGAGAACATGGTTTTCAAGTCGCACAAGAGGTATTC 360
Db     156 ThrAlaLeuAlaPheArgLeuLeuArgGlnTyrGlyPheLysValProGlnGluValPhe 175

Qy     361 GATAGTTTCAAGAACGAGGAGGGTGAGTTCAAAGAAAGCCTTAGCGACGACACCAGAGGA 420
Db     176 AspCysPheLysAsnAspAsnGlyGluPheLysArgSerLeuSerSerAspThrLysGly 195

Qy     421 TTGTTGCAACTGTATGAAGCTTCCTTTCTGTTGACGGAAGGCGAAACCACGCTCGAGTCA 480
Db     196 LeuLeuGlnLeuTyrGluAlaSerPheLeuLeuThrGluGlyGluMetThrLeuGluLeu 215

Qy     481 GCGAGGGAATTCGCCACCAAATTTTTGGAGGAAAAAGTGAACGAGGGTGGTGTTGATGGC 540
Db     216 AlaArgGluPheAlaThrIlePheLeuGlnGluLysLeuAsnAspLysThrIleAspAsp 235

Qy     541 GAC-----------------CTTTTAACAAGAATCGCATATTCTTTGGACATCCCTCTT 582
Db     236 AspAspAspAlaAspThrAsnLeuIleSerCysValArgHisSerLeuAspIleProIle 255

Qy     583 CATTGGAGGATTAAAAGGCCAAATGCACCTGTGTGGATCGAATGGTATAGGAAGAGGCCC 642
Db     256 HisTrpArgIleGlnArgProAsnAlaSerTrpTrpIleAspAlaTyrLysArgArgSer 275

Qy     643 GACATGAATCCAGTAGTGTTGGAGCTTGCCATACTCGACTTAAATATTGTTCAAGCACAA 702
Db     276 HisMetAsnProLeuValLeuGluLeuAlaLysLeuAspLeuAsnIlePheGlnAlaGln 295

Qy     703 TTTCAAGAAGAGCTCAAAGAATCCTTCAGGTGGTGGAGAAATACTGGGTTTGTTGAGAAG 762
Db     296 PheGlnGlnGluLeuLysGlnAspLeuGlyTrpTrpLysAsnThrCysLeuAlaGluLys 315

Qy     763 CTGCCCTTCGCAAGGGATAGACTGGTGGAATGCTACTTTTGGAATACTGGGATCATCGAG 822
Db     316 LeuProPheAlaArgAspArgLeuValGluCysTyrPheTrpCysThrGlyIleIleGln 335

Qy     823 CCACGTCAGCATGCAAGTGCAAGGATAATGATGGGCAAAGTCAACGCTCTGATTACGGTG 882
Db     336 ProLeuGlnHisGluAsnAlaArgValThrLeuAlaLysValAsnAlaLeuIleThrThr 355

Qy     883 ATCGATGATATTTATGATGTCTATGGCACCTTAGAAGAACTCGAACAATTCACTGACCTC 942
Db     356 LeuAspAspIleTyrAspValTyrGlyThrLeuGluGluLeuGluLeuPheThrGluAla 375

Qy     943 ATTCGAAGATGGGATATAAACTCAATCGACCAACTTCCCGATTACATGCAACTGTGCTTT 1002
Db     376 IleArgArgTrpAspValSerSerIleAspHisLeuProAsnTyrMetGlnLeuCysPhe 395

Qy    1003 CTTGCACTCAACAACTTCGTCGATGATACATCGTACGATGTTATGAAGGAGAAAGGCGTC 1062
Db     396 LeuAlaLeuAsnAsnPheValAspAspThrAlaTyrAspValMetLysGluLysAspIle 415

Qy    1063 AACGTTATACCCTACCTGCGGCAATCGTGGGTTGATTTGGCGGATAAGTATATGGTAGAG 1122
Db     416 AsnIleIleProTyrLeuArgLysSerTrpLeuAspLeuAlaGluThrTyrLeuValGlu 435

Qy    1123 GCACGGTGGTTCTACGGCGGGCACAAACCAAGTTTGGAAGAGTATTTGGAGAACTCATGG 1182
Db     436 AlaLysTrpPheTyrSerGlyHisLysProAsnMetGluGluTyrLeuAsnAsnAlaTrp 455

Qy    1183 CAGTCGATAAGTGGGCCCTGTATGTTAACGCACATATTCTTCCGAGTAACAGATTCGTTC 1242
```

```
                     ||||||||||||||    ||||||    |||:::||||||||||||||||||||
Db        456 IleSerIleSerGlyProValMetLeuCysHisValPhePheArgValThrAspSerIle 475

Qy       1243 ACAAAGGAGACCGTCGACAGTTTGTACAAATACCACGATTTAGTTCGTTGGTCATCCTTC 1302
              |||:::||||||:::||||||:::|||||||||||||||:::|||:::||||||
Db        476 ThrArgGluThrValGluSerLeuPheLysTyrHisAspLeuIleArgTyrSerSerThr 495

Qy       1303 GTTCTGCGGCTTGCTGATGATTTGGGAACCTCGGTGGAAGAGGTGAGCAGAGGGGATGTG 1362
              :::||||||||||||||||||||||||||||||:::||||||||||||||||||||||||
Db        496 IleLeuArgLeuAlaAspAspLeuGlyThrSerLeuGluGluValSerArgGlyAspVal 515

Qy       1363 CCGAAATCACTTCAGTGCTACATGAGTGACTACAATGCATCGGAGGCGGAGGCGCGGAAG 1422
              |||||||||:::|||||||||||||:::|||   ||||||||||||   |||||||||:::
Db        516 ProLysSerIleGlnCysTyrMetAsnAspAsnAsnAlaSerGluGluGluAlaArgArg 535

Qy       1423 CACGTGAAATGGCTGATAGCGGAGGTGTGGAAGAAGATGAATGCGGAGAGGGTGTCGAAG 1482
              ||||||:::||||||||||||||||   |||||||||:::|||    |||        |||
Db        536 HisValArgTrpLeuIleAlaGluThrTrpLysLysIleAsnGluGluValTrpSerAla 555

Qy       1483 GATTCTCCATTCGGCAAAGATTTTATAGGATGTGCAGTTGATTTAGGAAGGATGGCGCAG 1542
              ||||||||||||   ||||||||||||   ||||||   |||:::||||||||||||||||||
Db        556 AspSerProPheCysLysAspPheIleAlaCysAlaAlaAspMetGlyArgMetAlaGln 575

Qy       1543 TTGATGTACCATAATGGAGATGGGCACGGCACACAACACCCTATTATACATCAACAAATG 1602
               ||||||||||||||||||||||||||||||   |||:::|||   |||||||||||||||||
Db        576 PheMetTyrHisAsnGlyAspGlyHisGlyIleGlnAsnProGlnIleHisGlnGlnMet 595

Qy       1603 ACCAGAACCTTATTCGAG 1620
              |||     |||||||||
Db        596 ThrAspIleLeuPheGlu 601
```

```
RESULT 8
O04806
ID   O04806     PRELIMINARY;     PRT;   603 AA.
AC   O04806;
DT   01-JUL-1997 (TrEMBLrel. 04, Created)
DT   01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   LIMONENE cyclase.
GN   GPFLC.
OS   Perilla frutescens.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC   lamiids; Lamiales; Lamiaceae; Nepetoideae; Elsholtzieae; Perilla.
OX   NCBI_TaxID=48386;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NO.9;
RX   MEDLINE=96400360; PubMed=8806736;
RA   Yuba A., Yazaki K., Tabata M., Honda G., Croteau R.;
RT   "cDNA cloning, characterization, and functional expression of 4S-(-)-
RT   limonene synthase from Perilla frutescens.";
RL   Arch. Biochem. Biophys. 332:280-287(1996).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Tsubouchi T., Hara M., Yazaki K., Honda G.;
RL   Submitted (JUL-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; D49368; BAA08367.1; -.
DR   EMBL; AB005744; BAA21629.1; -.
DR   HSSP; Q40577; 5EAU.
DR   GO; GO:0016829; F:lyase activity; IEA.
DR   GO; GO:0008152; P:metabolism; IEA.
DR   InterPro; IPR005630; Terpene_synth_C.
DR   InterPro; IPR008949; Terpenoid_synth.
DR   InterPro; IPR008930; Terp_cyc_toroid.
DR   InterPro; IPR001906; Terp_synth-like.
DR   Pfam; PF01397; Terpene_synth; 1.
DR   Pfam; PF03936; Terpene_synth_C; 1.
SQ   SEQUENCE   603 AA;  70321 MW;  67F41F0C6C871181 CRC64;

Alignment Scores:
Pred. No.:            3.61e-149     Length:        603
Score:                2003.50       Matches:       375
Percent Similarity:   81.47%        Conservative:  69
Best Local Similarity: 68.81%       Mismatches:    90
Query Match:          68.90%        Indels:        11
DB:                   10            Gaps:          2

US-09-938-956-6 (1-1632) x O04806 (1-603)
```

```
Qy          4 AGACGATCCGGAAACTACAACCCTTCTCGTTGGGATGTCAACTTCATCCAATCGCTTCTC 63
              ||||||||||||||||||:::||||||   |||:::    :::::|||   ||||||
Db         61 ArgArgSerGlyAsnTyrSerProSerPheTrpAsnAlaAspTyrIleLeuSerLeuAsn 80

Qy         64 AGTGACTATAAGGAGGACAAACACGTGATTAGGGCTTCTGAGCTGGTCACTTTGGTGAAG 123
              |||   ||||||:::      |||:::   ||||||   ||||||:::      ||||||
Db         81 SerHisTyrLysAspLysSerHisMetLysArgAlaGlyGluLeuIleValGlnValLys 100

Qy        124 ATGGAACTGGAGAAAGAAACGGATCAAATTCGACAACTTGAGTTGATCGATGACTTGCAG 183
              |||   :::   ||||||||||||   :::   ||||||||||||||||||||||||||||
Db        101 MetValMetGlyLysGluThrAspProValValGlnLeuGluLeuIleAspAspLeuGln 120

Qy        184 AGGATGGGGCTGTCCGATCATTTCCAAAATGAGTTCAAAGAAATCTTGTCCTCTATATAT 243
              ::::::   ||||||   |||   :::   |||   ||||||||||||   |||
Db        121 LysLeuAlaLeuSerHisHisValGluLysGluIleLysGluIleLeuPheLysIleSer 140

Qy        244 CTCGACCATCACTATTACAAGAACCCTTTTCCAAAAGAAGAAAGGGATCTCTACTCCACA 303
                      |||                  ||||||||||||||||||||||
Db        141 ThrTyrAspHisLysIle---------------MetValGluArgAspLeuTyrSerThr 155

Qy        304 TCTCTTGCATTTAGGCTCCTCAGAGAACATGGTTTTCAAGTCGCACAAGAGGTATTCGAT 363
              :::|||||||||||||||||||||||::::::||||||:::|||   |||||||||||||||
Db        156 AlaLeuAlaPheArgLeuLeuArgGlnTyrGlyPheLysValProGlnGluValPheAsp 175

Qy        364 AGTTTCAAGAACGAGGAGGGTGAGTTCAAAGAAAGCCTTAGCGACGACACCAGAGGATTG 423
                ||||||||:::   ||||||||||||   ||||||   ||||||:::||||
Db        176 CysPheLysAsnAspAsnGlyGluPheLysArgSerLeuSerSerAspThrLysGlyLeu 195

Qy        424 TTGCAACTGTATGAAGCTTCCTTTCTGTTGACGGAAGGCGAAACCACGCTCGAGTCAGCG 483
              ||||||||||||||||||||||||||||||||||||   |||||||||   |||
Db        196 LeuGlnLeuTyrGluAlaSerPheLeuLeuThrGluGlyGluMetThrLeuGluLeuAla 215

Qy        484 AGGGAATTCGCCACCAAATTTTTGGAGGAAAAAGTGAACGAGGGTGGTGTTGATGGCGAC 543
              ||||||||||||||||||   |||:::||||||:::||||||      :::|||   |||
Db        216 ArgGluPheAlaThrLysSerLeuGlnGluLysLeuAsnGluLysThrIleAspAspAsp 235

Qy        544 ------------------CTTTTAACAAGAATCGCATATTCTTTGGACATCCCTCTTCAT 585
                                |||::::::   :::   :::||||||||||||:::||||||
Db        236 AspAspAlaAspThrAsnLeuIleSerCysValArgHisSerLeuAspIleProIleHis 255

Qy        586 TGGAGGATTAAAAGGCCAAATGCACCTGTGTGGATCGAATGGTATAGGAAGAGGCCCGAC 645
              ||||||||:::||||||||||||         ||||||:::   |||::::::|||
Db        256 TrpArgIleGlnArgProAsnAlaSerTrpTrpIleAspAlaTyrLysArgArgSerHis 275

Qy        646 ATGAATCCAGTAGTGTTGGAGCTTGCCATACTCGACTTAAATATTGTTCAAGCACAATTT 705
              |||||||||:::|||||||||||||||   |||||||||||||||   ||||||||||||
Db        276 MetAsnProLeuValLeuGluLeuAlaLysLeuAspLeuAsnIlePheGlnAlaGlnPhe 295

Qy        706 CAAGAAGAGCTCAAAGAATCCTTCAGGTGGTGGAGAAATACTGGGTTTGTTGAGAAGCTG 765
              |||:::|||||||||:::          |||||||:::||||||      |||||||||
Db        296 GlnGlnGluLeuLysGlnAspLeuGlyTrpTrpLysAsnThrCysLeuAlaGluLysLeu 315

Qy        766 CCCTTCGCAAGGGATAGACTGGTGGAATGCTACTTTTGGAATACTGGGATCATCGAGCCA 825
              ||||||   |||||||||||||||||||||||||||||   ||||||||||||:::|||
Db        316 ProPheValArgAspArgLeuValGluCysTyrPheTrpCysThrGlyIleIleGlnPro 335

Qy        826 CGTCAGCATGCAAGTGCAAGGATAATGATGGGCAAAGTCAACGCTCTGATTACGGTGATC 885
                 ||||||   :::|||||||:::   :::   |||||||||||||||||||||   :::
Db        336 LeuGlnHisGluAsnAlaArgValThrLeuAlaLysValAsnAlaLeuIleThrThrLeu 355

Qy        886 GATGATATTTATGATGTCTATGGCACCTTAGAAGAACTCGAACAATTCACTGACCTCATT 945
              ||||||||||||||||||||||||||||||||||||||||||||   ||||||:::   |||
Db        356 AspAspIleTyrAspValTyrGlyThrLeuGluGluLeuGluLeuPheThrGluAlaIle 375

Qy        946 CGAAGATGGGATATAAACTCAATCGACCAACTTCCCGATTACATGCAACTGTGCTTTCTT 1005
              ||||||||||||::::::|||||||||   ||||||:::||||||||||||||||||||
```

```
Db        376 ArgArgTrpAspValSerSerIleAspHisLeuProAsnTyrMetGlnLeuCysPheLeu 395

Qy       1006 GCACTCAACAACTTCGTCGATGATACATCGTACGATGTTATGAAGGAGAAAGGCGTCAAC 1065

Db        396 AlaLeuAsnAsnPheValAspAspThrAlaTyrAspValMetLysGluLysAspIleAsn 415

Qy       1066 GTTATACCCTACCTGCGGCAATCGTGGGTTGATTTGGCGGATAAGTATATGGTAGAGGCA 1125

Db        416 IleIleProTyrLeuArgLysSerTrpLeuAspLeuAlaGluThrTyrLeuValGluAla 435

Qy       1126 CGGTGGTTCTACGGCGGGCACAAACCAAGTTTGGAAGAGTATTTGGAGAACTCATGGCAG 1185

Db        436 LysTrpPheTyrSerGlyHisLysProAsnLeuGluGluTyrLeuAsnAsnAlaTrpIle 455

Qy       1186 TCGATAAGTGGGCCCTGTATGTTAACGCACATATTCTTCCGAGTAACAGATTCGTTCACA 1245

Db        456 SerIleSerGlyProValMetLeuTrpHisValPhePheArgValThrAspSerIleThr 475

Qy       1246 AAGGAGACCGTCGACAGTTTGTACAAATACCACGATTTAGTTCGTTGGTCATCCTTCGTT 1305

Db        476 ArgGluThrValGluSerLeuPheLysTyrHisAspLeuIleArgTyrSerSerThrIle 495

Qy       1306 CTGCGGCTTGCTGATGATTTGGGAACCTCGGTGGAAGAGGTGAGCAGAGGGGATGTGCCG 1365

Db        496 LeuArgLeuAlaAspAspLeuGlyThrSerLeuGluGluValSerArgGlyAspValPro 515

Qy       1366 AAATCACTTCAGTGCTACATGAGTGACTACAATGCATCGGAGGCGGAGGCGCGGAAGCAC 1425

Db        516 LysSerIleGlnCysTyrMetAsnAspAsnAsnAlaSerGluGluGluAlaArgArgHis 535

Qy       1426 GTGAAATGGCTGATAGCGGAGGTGTGGAAGAAGATGAATGCGGAGAGGGTGTCGAAGGAT 1485

Db        536 IleArgTrpLeuIleAlaGluThrTrpLysLysIleAsnGluGluValTrpSerValAsp 555

Qy       1486 TCTCCATTCGGCAAAGATTTTATAGGATGTGCAGTTGATTTAGGAAGGATGGCGCAGTTG 1545

Db        556 SerProPheCysLysAspPheIleAlaCysAlaAlaAspMetGlyArgMetAlaGlnPhe 575

Qy       1546 ATGTACCATAATGGAGATGGGCACGGCACACAACACCCTATTATACATCAACAAATGACC 1605

Db        576 MetTyrHisAsnGlyAspGlyHisGlyIleGlnAsnProGlnIleHisGlnGlnMetThr 595

Qy       1606 AGAACCTTATTCGAG 1620

Db        596 AspIleLeuPheGlu 600

RESULT 9
Q9AXM7
ID   Q9AXM7      PRELIMINARY;      PRT;   604 AA.
AC   Q9AXM7;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Limonene synthase.
OS   Perilla frutescens var. frutescens.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC   lamiids; Lamiales; Lamiaceae; Nepetoideae; Elsholtzieae; Perilla.
OX   NCBI_TaxID=151328;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=5520;
RA   Takeya H., Ito M., Honda G.;
RT   "Limonene synthase from PT type Perilla frutescens.";
RL   Submitted (OCT-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF317695; AAK06663.1; -.
DR   HSSP; Q40577; 5EAT.
DR   GO; GO:0016829; F:lyase activity; IEA.
DR   GO; GO:0008152; P:metabolism; IEA.
DR   InterPro; IPR005630; Terpene_synth_C.
DR   InterPro; IPR008949; Terpenoid_synth.
DR   InterPro; IPR008930; Terp_cyc_toroid.
DR   InterPro; IPR001906; Terp_synth-like.
DR   Pfam; PF01397; Terpene_synth; 1.
DR   Pfam; PF03936; Terpene_synth_C; 1.
SQ   SEQUENCE   604 AA;  70528 MW;  7A8A8B3821774B59 CRC64;

Alignment Scores:
Pred. No.:               4.74e-149     Length:          604
Score:                   2002.00       Matches:         374
Percent Similarity:      82.05%        Conservative:    74
Best Local Similarity:   68.50%        Mismatches:      86
Query Match:             68.84%        Indels:          12
DB:                      10            Gaps:            3

US-09-938-956-6 (1-1632) x Q9AXM7 (1-604)

Qy          4 AGACGATCCGGAAACTACAACCCTTCTCGTTGGGATGTCAACTTCATCCAATCGCTTCTC 63

Db         61 ArgArgSerGlyAsnTyrSerProSerPheTrpAsnAlaAspTyrIleLeuSerLeuAsn 80

Qy         64 AGTGACTATAAGGAGGAC---AAACACGTGATTAGGGCTTCTGAGCTGGTCACTTTGGTG 120

Db         81 AsnHisTyrLysGluGluSerArgHisMetLysArgAlaGlyGluLeuIleValGlnVal 100

Qy        121 AAGATGGAACTGGAGAAAGAAACGGATCAAATTCGACAACTTGAGTTGATCGATGACTTG 180

Db        101 LysMetValMetGlyLysGluThrAspProValValGlnLeuGluLeuIleAspAspLeu 120

Qy        181 CAGAGGATGGGGCTGTCCGATCATTTCCAAAATGAGTTCAAAGAAATCTTGTCCTCTATA 240

Db        121 HisLysLeuAlaLeuSerHisHisPheGluLysGluIleLysGluIleLeuPheAsnIle 140

Qy        241 TATCTCGACCATCACTATTACAAGAACCCTTTTCCAAAAGAAGAAAGGGATCTCTACTCC 300

Db        141 SerIleTyrAspHisLysIle---------------MetValGluArgAspLeuTyrSer 155

Qy        301 ACATCTCTTGCATTTAGGCTCCTCAGAGAACATGGTTTTCAAGTCGCACAAGAGGTATTC 360

Db        156 ThrAlaLeuAlaPheArgLeuLeuArgGlnTyrGlyPheLysValProGlnGluValPhe 175

Qy        361 GATAGTTTCAAGAACGAGGAGGGTGAGTTCAAAGAAAGCCTTAGCGACGACACCAGAGGA 420

Db        176 AspCysPheLysAsnAspAsnGlyGluPheLysArgSerLeuSerSerAspThrLysGly 195

Qy        421 TTGTTGCAACTGTATGAAGCTTCCTTTCTGTTGACGGAAGGCGAAACCACGCTCGAGTCA 480

Db        196 LeuLeuGlnLeuTyrGluAlaSerPheLeuLeuThrGluGlyGluMetThrLeuGluLeu 215

Qy        481 GCGAGGGAATTCGCCACCAAATTTTTGGAGGAAAAAGTGAACGAGGGTGGTGTTGATGGC 540

Db        216 AlaArgGluPheAlaThrIlePheLeuGlnGluLysLeuAsnAspLysThrIleAspAsp 235

Qy        541 GAC-----------------CTTTTAACAAGAATCGCATATTCTTTGGACATCCCTCTT 582

Db        236 AspAspAspAlaAspThrAsnLeuIleSerCysValArgHisSerLeuAspIleProIle 255

Qy        583 CATTGGAGGATTAAAAGGCCAAATGCACCTGTGTGGATCGAATGGTATAGGAAGAGGCCC 642

Db        256 HisTrpArgIleGlnArgProAsnAlaSerTrpTrpIleAspAlaTyrLysArgArgSer 275

Qy        643 GACATGAATCCAGTAGTGTTGGAGCTTGCCATACTCGACTTAAATATTGTTCAAGCACAA 702

Db        276 HisMetAsnProLeuValLeuGluLeuAlaLysLeuAspLeuAsnIlePheGlnAlaGln 295

Qy        703 TTTCAAGAAGAGCTCAAAGAATCCTTCAGGTGGTGGAGAAATACTGGGTTTGTTGAGAAG 762

Db        296 PheGlnGlnGluLeuLysGlnAspLeuGlyTrpTrpLysAsnThrCysLeuAlaGluLys 315

Qy        763 CTGCCCTTCGCAAGGGATAGACTGGTGGAATGCTACTTTTGGAATACTGGGATCATCGAG 822

Db        316 LeuProPheThrArgAspArgLeuValGluCysTyrPheTrpCysThrGlyIleIleGln 335

Qy        823 CCACGTCAGCATGCAAGTGCAAGGATAATGATGGGCAAAGTCAACGCTCTGATTACGGTG 882

Db        336 ProLeuGlnHisGluAsnAlaArgValThrLeuAlaLysValAsnAlaLeuIleThrThr 355
```

```
Qy        883 ATCGATGATATTTATGATGTCTATGGCACCTTAGAAGAACTCGAACAATTCACTGACCTC 942
Db        356 LeuAspAspIleTyrAspValTyrGlyThrLeuGluGluLeuGluLeuPheThrGluAla 375

Qy        943 ATTCGAAGATGGGATATAAACTCAATCGACCAACTTCCCGATTACATGCAACTGTGCTTT 1002
Db        376 IleArgArgTrpAspValSerSerIleAspHisLeuProAsnTyrMetGlnLeuCysPhe 395

Qy       1003 CTTGCACTCAACAACTTCGTCGATGATACATCGTACGATGTTATGAAGGAGAAAGGCGTC 1062
Db        396 LeuAlaLeuAsnAsnPheValAspAspThrAlaTyrAspValMetLysGluLysAspIle 415

Qy       1063 AACGTTATACCCTACCTGCGGCAATCGTGGGTTGATTTGGCGGATAAGTATATGGTAGAG 1122
Db        416 AsnIleIleProTyrLeuArgLysSerTrpLeuAspLeuAlaGluThrTyrLeuValGlu 435

Qy       1123 GCACGGTGGTTCTACGGCGGGCACAAACCAAGTTTGGAAGAGTATTTGGAGAACTCATGG 1182
Db        436 AlaLysTrpPheTyrSerGlyHisLysProAsnMetGluGluTyrLeuAsnAsnAlaTrp 455

Qy       1183 CAGTCGATAAGTGGGCCCTGTATGTTAACGCACATATTCTTCCGAGTAACAGATTCGTTC 1242
Db        456 IleSerIleSerGlyProValMetLeuCysHisValPhePheArgValThrAspSerIle 475

Qy       1243 ACAAAGGAGACCGTCGACAGTTTGTACAAATACCACGATTTAGTTCGTTGGTCATCCTTC 1302
Db        476 ThrArgGluThrValGluSerLeuPheLysTyrHisAspLeuIleArgTyrSerSerThr 495

Qy       1303 GTTCTGCGGCTTGCTGATGATTTGGGAACCTCGGTGGAAGAGGTGAGCAGAGGGGATGTG 1362
Db        496 IleLeuArgLeuAlaAspAspLeuGlyThrSerLeuGluGluValSerArgGlyAspVal 515

Qy       1363 CCGAAATCACTTCAGTGCTACATGAGTGACTACAATGCATCGGAGGCGGAGGCGCGGAAG 1422
Db        516 ProLysSerIleGlnCysTyrMetAsnAspAsnAsnAlaSerGluGluGluAlaArgArg 535

Qy       1423 CACGTGAAATGGCTGATAGCGGAGGTGTGGAAGAAGATGAATGCGGAGAGGGTGTCGAAG 1482
Db        536 HisValArgTrpLeuIleAlaGluThrTrpLysLysIleAsnGluGluValTrpSerAla 555

Qy       1483 GATTCTCCATTCGGCAAAGATTTTATAGGATGTGCAGTTGATTTAGGAAGGATGGCGCAG 1542
Db        556 AspSerProPheCysLysAspPheIleAlaCysAlaAlaAspMetGlyArgMetAlaGln 575

Qy       1543 TTGATGTACCATAATGGAGATGGGCACGGCACACAACACCCTATTATACATCAACAAATG 1602
Db        576 PheMetTyrHisAsnGlyAspGlyHisGlyIleGlnAsnProGlnIleHisGlnGlnMet 595

Qy       1603 ACCAGAACCTTATTCGAG 1620
Db        596 ThrAspIleLeuPheGlu 601

RESULT 10
Q9FV72
ID   Q9FV72      PRELIMINARY;      PRT;   603 AA.
AC   Q9FV72;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Limonene synthase.
OS   Perilla frutescens.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC   lamiids; Lamiales; Lamiaceae; Nepetoideae; Elsholtzieae; Perilla.
OX   NCBI_TaxID=48386;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=1;
RA   Dodo R., Ito M., Takeya H., Honda G.;
RT   "Limonene synthase from Perilla sp.";
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF241793; AAG31438.1; -.
DR   HSSP; Q40577; 5EAU.
DR   GO; GO:0016829; F:lyase activity; IEA.
DR   GO; GO:0008152; P:metabolism; IEA.
DR   InterPro; IPR005630; Terpene_synth_C.
DR   InterPro; IPR008949; Terpenoid_synth.
DR   InterPro; IPR008930; Terp_cyc_toroid.
DR   InterPro; IPR001906; Terp_synth-like.
DR   Pfam; PF01397; Terpene_synth; 1.
DR   Pfam; PF03936; Terpene_synth_C; 1.
SQ   SEQUENCE   603 AA;  70264 MW;  653EC5C7F2520AC8 CRC64;

Alignment Scores:
Pred. No.:              8.93e-149      Length:          603
Score:                  1998.50        Matches:         374
Percent Similarity:     81.28%         Conservative:    69
Best Local Similarity:  68.62%         Mismatches:      91
Query Match:            68.72%         Indels:          11
DB:                     10             Gaps:            2

US-09-938-956-6 (1-1632) x Q9FV72 (1-603)

Qy          4 AGACGATCCGGAAACTACAACCCTTCTCGTTGGGATGTCAACTTCATCCAATCGCTTCTC 63
Db         61 ArgArgSerGlyAsnTyrSerProSerPheTrpAsnAlaAspTyrIleLeuLeuLeuAsn 80

Qy         64 AGTGACTATAAGGAGGACAAACACGTGATTAGGGCTTCTGAGCTGGTCACTTTGGTGAAG 123
Db         81 SerHisTyrLysAspLysSerHisMetLysArgAlaGlyGluLeuIleValGlnValLys 100

Qy        124 ATGGAACTGGAGAAAGAAACGGATCAAATTCGACAACTTGAGTTGATCGATGACTTGCAG 183
Db        101 MetValMetGlyLysGluThrAspProValValGlnLeuGluLeuIleAspAspLeuGln 120

Qy        184 AGGATGGGGCTGTCCGATCATTTCCAAAATGAGTTCAAAGAAATCTTGTCCTCTATATAT 243
Db        121 LysLeuAlaLeuSerHisHisValGluLysGluIleLysGluIleLeuPheLysIleSer 140

Qy        244 CTCGACCATCACTATTACAAGAACCCTTTTCCAAAAGAAGAAAGGGATCTCTACTCCACA 303
Db        141 ThrTyrAspHisLysIle---------------MetValGluArgAspLeuTyrSerThr 155

Qy        304 TCTCTTGCATTTAGGCTCCTCAGAGAACATGGTTTTCAAGTCGCACAAGAGGTATTCGAT 363
Db        156 AlaLeuAlaPheArgLeuLeuArgGlnTyrGlyPheLysValProGlnGluValPheAsp 175

Qy        364 AGTTTCAAGAACGAGGAGGGTGAGTTCAAAGAAAGCCTTAGCGACGACACCAGAGGATTG 423
Db        176 CysPheLysAsnAspAsnGlyGluPheLysArgSerLeuSerSerAspThrLysGlyLeu 195

Qy        424 TTGCAACTGTATGAAGCTTCCTTTCTGTTGACGGAAGGCGAAACCACGCTCGAGTCAGCG 483
Db        196 LeuGlnLeuTyrGluAlaSerPheLeuLeuThrGluGlyGluMetThrLeuGluLeuAla 215

Qy        484 AGGGAATTCGCCACCAAATTTTTGGAGGAAAAAGTGAACGAGGGTGGTGTTGATGGCGAC 543
Db        216 ArgGluPheAlaThrLysSerLeuGlnGluLysLeuAsnGluLysThrIleAspAspAsp 235

Qy        544 -----------------CTTTTAACAAGAATCGCATATTCTTTGGACATCCCTCTTCAT 585
Db        236 AspAspAlaAspThrAsnLeuIleSerCysValArgHisSerLeuAspIleProIleHis 255

Qy        586 TGGAGGATTAAAAGGCCAAATGCACCTGTGTGGATCGAATGGTATAGGAAGAGGCCCGAC 645
Db        256 TrpArgIleGlnArgProAsnAlaSerTrpTrpIleAspAlaTyrLysArgArgSerHis 275

Qy        646 ATGAATCCAGTAGTGTTGGAGCTTGCCATACTCGACTTAAATATTGTTCAAGCACAATTT 705
Db        276 MetAsnProLeuValLeuGluLeuAlaLysLeuAspLeuAsnIlePheGlnAlaGlnPhe 295

Qy        706 CAAGAAGAGCTCAAAGAATCCTTCAGGTGGTGGAGAAATACTGGGTTTGTTGAGAAGCTG 765
Db        296 GlnGlnGluLeuLysGlnAspLeuGlyTrpTrpLysAsnThrCysLeuAlaGluLysLeu 315
```

```
Qy        766 CCCTTCGCAAGGGATAGACTGGTGGAATGCTACTTTTGGAATACTGGGATCATCGAGCCA 825
Db        316 ProPheValArgAspArgLeuValGluCysTyrPheTrpCysThrGlyIleIleGlnPro 335

Qy        826 CGTCAGCATGCAAGTGCAAGGATAATGATGGGCAAAGTCAACGCTCTGATTACGGTGATC 885
Db        336 LeuGlnHisGluAsnAlaArgValThrLeuAlaLysValAsnAlaLeuIleThrThrLeu 355

Qy        886 GATGATATTTATGATGTCTATGGCACCTTAGAAGAACTCGAACAATTCACTGACCTCATT 945
Db        356 AspAspIleTyrAspValTyrGlyThrLeuGluGluLeuGluLeuPheThrGluAlaIle 375

Qy        946 CGAAGATGGGATATAAACTCAATCGACCAACTTCCCGATTACATGCAACTGTGCTTTCTT 1005
Db        376 ArgArgTrpAspValSerSerIleAspHisLeuProAsnTyrMetGlnLeuCysPheLeu 395

Qy       1006 GCACTCAACAACTTCGTCGATGATACATCGTACGATGTTATGAAGGAGAAAGGCGTCAAC 1065
Db        396 AlaLeuAsnAsnPheValAspAspThrAlaTyrAspValMetLysGluLysAspIleAsn 415

Qy       1066 GTTATACCCTACCTGCGGCAATCGTGGGTTGATTTGGCGGATAAGTATATGGTAGAGGCA 1125
Db        416 IleIleProTyrLeuArgLysSerTrpLeuAspLeuAlaGluThrTyrLeuValGluAla 435

Qy       1126 CGGTGGTTCTACGGCGGGCACAAACCAAGTTTGGAAGAGTATTTGGAGAACTCATGGCAG 1185
Db        436 LysTrpPheTyrSerGlyHisLysProAsnLeuGluGluTyrLeuAsnAsnAlaTrpIle 455

Qy       1186 TCGATAAGTGGGCCCTGTATGTTAACGCACATATTCTTCCGAGTAACAGATTCGTTCACA 1245
Db        456 SerIleSerGlyProValMetLeuCysHisValPhePheArgValThrAspSerIleThr 475

Qy       1246 AAGGAGACCGTCGACAGTTTGTACAAATACCACGATTTAGTTCGTTGGTCATCCTTCGTT 1305
Db        476 ArgGluThrValGluSerLeuPheLysTyrHisAspLeuIleArgTyrSerSerThrIle 495

Qy       1306 CTGCGGCTTGCTGATGATTTGGGAACCTCGGTGGAAGAGGTGAGCAGAGGGGATGTGCCG 1365
Db        496 LeuArgLeuAlaAspAspLeuGlyThrSerLeuGluGluValSerArgGlyAspValPro 515

Qy       1366 AAATCACTTCAGTGCTACATGAGTGACTACAATGCATCGGAGGCGGAGGCGCGGAAGCAC 1425
Db        516 LysSerIleGlnCysTyrMetAsnAspAsnAsnAlaSerGluGluGluAlaArgArgHis 535

Qy       1426 GTGAAATGGCTGATAGCGGAGGTGTGGAAGAAGATGAATGCGGAGAGGGTGTCGAAGGAT 1485
Db        536 IleArgTrpLeuIleAlaGluThrTrpLysLysIleAsnGluGluValTrpSerValAsp 555

Qy       1486 TCTCCATTCGGCAAAGATTTTATAGGATGTGCAGTTGATTTAGGAAGGATGGCGCAGTTG 1545
Db        556 SerProPheCysLysAspPheIleAlaCysAlaAlaAspMetGlyArgMetAlaGlnPhe 575

Qy       1546 ATGTACCATAATGGAGATGGCACGGCACACAACACCCTATTATACATCAACAAATGACC 1605
Db        576 MetTyrHisAsnGlyAspGlyHisGlyIleGlnAsnProGlnIleHisGlnGlnMetThr 595

Qy       1606 AGAACCTTATTCGAG 1620
Db        596 AspIleLeuPheGlu 600
```

RESULT 11
Q8L5J7

```
ID   Q8L5J7      PRELIMINARY;      PRT;    597 AA.
AC   Q8L5J7;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   3-carene synthase.
OS   Salvia stenophylla.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC   lamiids; Lamiales; Lamiaceae; Nepetoideae; Mentheae; Salvia.
OX   NCBI_TaxID=202612;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Glandular trichomes;
RA   Hoelscher D.J., Wildung M.R., Williams D.C., Croteau R.;
RT   "A cDNA clone for a 3-carene synthase from Salvia stenophylla.";
RL   Submitted (JUL-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF527416; AAM89254.1; -.
DR   GO; GO:0016829; F:lyase activity; IEA.
DR   GO; GO:0008152; P:metabolism; IEA.
DR   InterPro; IPR005630; Terpene_synth_C.
DR   InterPro; IPR008949; Terpenoid_synth.
DR   InterPro; IPR008930; Terp_cyc_toroid.
DR   InterPro; IPR001906; Terp_synth-like.
DR   Pfam; PF01397; Terpene_synth; 1.
DR   Pfam; PF03936; Terpene_synth_C; 1.
SQ   SEQUENCE   597 AA;  69714 MW;  4B562A8967B0F2F4 CRC64;
```

```
Alignment Scores:
Pred. No.:              5.46e-148    Length:         597
Score:                  1988.50      Matches:        367
Percent Similarity:     82.14%       Conservative:   79
Best Local Similarity:  67.59%       Mismatches:     94
Query Match:            68.38%       Indels:         3
DB:                     10           Gaps:           3
```

US-09-938-956-6 (1-1632) x Q8L5J7 (1-597)

```
Qy          4 AGACGATCCGGAAACTACAACCCTTCTCGTTGGGATGTCAACTTCATCCAATCGCTTCTC 63
Db         51 ArgArgSerGlyAsnTyrSerProSerLysTrpAspValAspTyrIleGlnSerLeuHis 70

Qy         64 AGTGACTATAAGGAGGACAAACACGTGATTAGGGCTTCTGAGCTGGTCACTTTGGTGAAG 123
Db         71 SerAspTyrLysGluGluArgHisThrArgArgAlaSerGluLeuIleMetGluValLys 90

Qy        124 ATGGAACTGGAGAAAGAAACGGATCAAATTCGACAACTTGAGTTGATCGATGACTTGCAG 183
Db         91 LysLeuLeuGluLysGluProAsnProThrArgGlnLeuGluLeuIleAspAspLeuGln 110

Qy        184 AGGATGGGGCTGTCCGATCATTTCCAAAATGAGTTCAAAGAAATCTTGTCCTCTATATAT 243
Db        111 LysLeuGlyLeuSerAspHisPheAsnAsnGluPheLysGluIleLeuAsnSerValTyr 130

Qy        244 CTCGACCATCACTATTACAAGAACCCTTTTCCAAAAGAA---GAAAGGGATCTCTACTCC 300
Db        131 LeuAspAsnLysTyrTyrArgAsnGlyAlaMetLysGluValGluArgAspLeuTyrSer 150

Qy        301 ACATCTCTTGCATTTAGGCTCCTCAGAGAACATGGTTTTCAAGTCGCACAAGAGGTATTC 360
Db        151 ThrAlaLeuAlaPheArgLeuLeuArgGlnHisGlyPheGlnValAlaGlnAspValLeu 170

Qy        361 GATAGTTTCAAGAACGAGGAGGGTGAGTTCAAAGAAAGCCTTAGCGACGACACCAGAGGA 420
Db        171 GluCysPheLysAsnThrLysGlyGluPheGluProSerLeuSerAspAspThrArgGly 190

Qy        421 TTGTTGCAACTGTATGAAGCTTCCTTTCTGTTGACGGAAGGCGAAACCACGCTCGAGTCA 480
Db        191 LeuLeuGlnLeuTyrGluAlaSerPheLeuLeuThrGluGlyGluAsnThrLeuGluLeu 210

Qy        481 GCGAGGGAATTCGCCACCAAATTTTTGGAGGAAAAAGTGAACGAGGGTGGTGTTGATGGC 540
Db        211 AlaArgAspPheThrThrLysIleLeuGluGluLysLeuArgAsnAspGluIleAspAsp 230

Qy        541 ---GACCTTTTAACAAGAATCGCATATTCTTTGGACATCCCTCTTCATTGGAGGATTAAA 597
Db        231 IleAsnLeuValThrTrpIleArgHisSerLeuGluIleProIleHisTrpArgIleAsp 250

Qy        598 AGGCCAAATGCACCTGTGTGGATCGAATGGTATAGGAAGAGGCCCGACATGAATCCAGTA 657
Db        251 ArgValAsnThrSerValTrpIleAspValTyrLysArgArgProAspMetAsnProIle 270

Qy        658 GTGTTGGAGCTTGCCATACTCGACTTAAATATTGTTCAAGCACAATTTCAAGAAGAGCTC 717
```

```
Db        271 ValLeuGluLeuAlaValLeuAspSerAsnIleValGlnAlaGlnTyrGlnGluGluLeu 290

Qy        718 AAAGAATCCTTCAGGTGGTGGAGAAATACTGGGTTTGTTGAGAAGCTGCCCTTCGCAAGG 777
Db        291 LysLeuAspLeuGlnTrpTrpArgAsnThrCysLeuAlaGluLysLeuProPheAlaArg 310

Qy        778 GATAGACTGGTGGAATGCTACTTTTGGAATACTGGGATCATCGAGCCACGTCAGCATGCA 837
Db        311 AspArgLeuValGluSerTyrPheTrpGlyValGlyValValGlnProArgGlnHisGly 330

Qy        838 AGTGCAAGGATAATGATGGGCAAAGTCAACGCTCTGATTACGGTGATCGATGATATTTAT 897
Db        331 IleAlaArgMetAlaValAspArgSerIleAlaLeuIleThrValIleAspAspValTyr 350

Qy        898 GATGTCTATGGCACCTTAGAAGAACTCGAACAATTCACTGACCTCATTCGAAGATGGGAT 957
Db        351 AspValTyrGlyThrLeuGluGluLeuGluGlnPheThrGluAlaIleArgArgTrpAsp 370

Qy        958 ATAAACTCAATCGACCAACTTCCCGATTACATGCAACTGTGCTTTCTTGCACTCAACAAC 1017
Db        371 IleSerSerIleAspGlnLeuProSerTyrMetGlnLeuCysPheLeuAlaLeuAspAsn 390

Qy       1018 TTCGTCGATGATACATCGTACGATGTTATGAAGGAGAAAGGCGTCAACGTTATACCCTAC 1077
Db        391 PheIleAsnAspIleAlaTyrAspValLeuLysGluGlnGlyPheAsnIleIleProTyr 410

Qy       1078 CTGCGGCAATCGTGGGTTGATTTGGCGGATAAGTATATGGTAGAGGCACGGTGGTTCTAC 1137
Db        411 LeuArgLysSerTrpThrAspMetIleGluGlyPheLeuLeuGluAlaLysTrpTyrHis 430

Qy       1138 GGCGGGCACAAACCAAGTTTGGAAGAGTATTTGGAGAACTCATGGCAGTCGATAAGTGGG 1197
Db        431 AsnGlyHisLysProLysLeuGluGluTyrLeuGluAsnGlyTrpArgSerIleGlySer 450

Qy       1198 CCCTGTATGTTAACGCACATATTCTTCCGAGTAACAGATTCGTTCACAAAGGAGACCGTC 1257
Db        451 ThrValValLeuThrHisAlaPhePheGlyValThrHisSerLeuThrLysGluAsnIle 470

Qy       1258 GACAGTTTGTACAAATACCACGATTTAGTTCGTTGGTCATCCTTCGTTCTGCGGCTTGCT 1317
Db        471 AspGlnPhePheGlyTyrHisGluIleValArgLeuSerSerMetLeuLeuArgLeuAla 490

Qy       1318 GATGATTTGGGAACCTCGGTGGAAGAGGTGAGCAGAGGGGATGTGCCGAAATCACTTCAG 1377
Db        491 AspAspLeuGlyThrSerThrAspGluValSerArgGlyAspValProLysAlaIleGln 510

Qy       1378 TGCTACATGAGTGAC---TACAATGCATCGGAGGCGGAGGCGCGGAAGCACGTGAAATGG 1434
Db        511 CysTyrMetAsnAspAsnIleGlyAlaSerGluAlaGluAlaArgGluHisValLysTrp 530

Qy       1435 CTGATAGCGGAGGTGTGGAAGAAGATGAATGCGGAGAGGGTGTCGAAGGATTCTCCATTC 1494
Db        531 CysIleTrpGluThrTrpLysLysMetAsnLysValArgValAlaArgAspThrProPhe 550

Qy       1495 GGCAAAGATTTTATAGGATGTGCAGTTGATTTAGGAAGGATGGCGCAGTTGATGTACCAT 1554
Db        551 SerGlnAspPheIleValCysAlaMetGlyMetGlyArgMetGlyGlnTyrMetTyrHis 570

Qy       1555 AATGGAGATGGGCACGGCACACAACACCCTATTATACATCAACAAATGACCAGAACCTTA 1614
Db        571 TyrGlyAspGlyHisGlyIleGlnHisSerIleIleHisGlnGlnMetSerThrCysLeu 590

Qy       1615 TTCGAGCCC 1623
Db        591 PheHisPro 593
```

```
RESULT 12
Q9LKM5
ID   Q9LKM5      PRELIMINARY;      PRT;    608 AA.
AC   Q9LKM5;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Terpene synthase.
OS   Perilla frutescens.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC   lamiids; Lamiales; Lamiaceae; Nepetoideae; Elsholtzieae; Perilla.
OX   NCBI_TaxID=48386;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=5526;
RA   Yagura T., Ito M., Honda G.;
RT   "Unknown terpene synthase from Perilla frutescens.";
RL   Submitted (MAY-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF271259; AAF76186.1; -.
DR   HSSP; Q40577; 5EAU.
DR   GO; GO:0016829; F:lyase activity; IEA.
DR   GO; GO:0008152; P:metabolism; IEA.
DR   InterPro; IPR005630; Terpene_synth_C.
DR   InterPro; IPR008949; Terpenoid_synth.
DR   InterPro; IPR008930; Terp_cyc_toroid.
DR   InterPro; IPR001906; Terp_synth-like.
DR   Pfam; PF01397; Terpene_synth; 1.
DR   Pfam; PF03936; Terpene_synth_C; 1.
SQ   SEQUENCE   608 AA;  70836 MW;  6A4D49AD55276010 CRC64;

Alignment Scores:
Pred. No.:              5.72e-144      Length:          608
Score:                  1937.50        Matches:         368
Percent Similarity:     81.42%         Conservative:    79
Best Local Similarity:  67.03%         Mismatches:      87
Query Match:            66.63%         Indels:          15
DB:                     10             Gaps:            5

US-09-938-956-6 (1-1632) x Q9LKM5 (1-608)
```

```
Qy          4 AGACGATCCGGAAACTACAACCCTTCTCGTTGGGATGTCAACTTCATCCAATCGCTTCTC 63
Db         58 ArgArgSerGlyAsnTyrSerProSerPheTrpAsnThrAspTyrIleLeuSerLeuAsn 77

Qy         64 AGTGACTATAAGGAGGACAAACACGTGATT---AGGGCTTCTGAGCTGGTCACTTTGGTG 120
Db         78 CysAspTyrGluAspGluArgArgMetArgGlyAlaAlaGlyGluLeuValGluGlnVal 97

Qy        121 AAGATGGAACTGGAGAAAGAAACGGATCAAATTCGACAACTTGAGTTGATCGATGACTTG 180
Db         98 LysMetLeuMetGluLysGluThrAspProIleValGlnLeuGluLeuIleAspValLeu 117

Qy        181 CAGAGGATGGGGCTGTCCGATCATTTCCAAAATGAGTTCAAAGAAATCTTG--------- 231
Db        118 GlnLysLeuAlaLeuSerHisHisPheGluLysGluPheGluGlyIleLeuPheAsnIle 137

Qy        232 TCCTCTATATATCTCGACCATCACTATTACAAGAACCCTTTTCCAAAAGAAGAAAGGGAT 291
Db        138 SerThrIleTyrAspAsp------------LysAsn------------ArgGluArgAsp 149

Qy        292 CTCTACTCCACATCTCTTGCATTTAGGCTCCTCAGAGAACATGGTTTTCAAGTCGCACAA 351
Db        150 LeuTyrSerThrThrLeuAlaPheArgLeuLeuArgGlnHisGlyTyrGlnValProGln 169

Qy        352 GAGGTATTCGATAGTTTCAAGAACGAGGAGGGTGAGTTCAAAGAAAGCCTTAGCGACGAC 411
Db        170 GluLeuPheGluCysPheLysAsnAspLysGlyGluPheLysGluSerLeuSerAsnAsp 189

Qy        412 ACCAGAGGATTGTTGCAACTGTATGAAGCTTCCTTTCTGTTGACGGAAGGCGAAACCACG 471
Db        190 ThrLysGlyLeuLeuGlnLeuTyrGluAlaSerPheLeuLeuThrGluGlyGluThrThr 209

Qy        472 CTCGAGTCAGCGAGGGAATTCGCCACCAAATTTTTGGAGGAAAAAGTGAAC--------- 522
Db        210 LeuGluLeuAlaArgGluPheAlaThrLysPheLeuGlnGluLysGluLysLeuAsnIle 229

Qy        523 GAGGGTGGTGTTGATGGCGACCTTTTAACAAGAATCGCATATTCTTTGGACATCCCTCTT 582
Db        230 AspAspAspAspAspThrAsnLeuIleSerCysValArgHisSerLeuAspMetProIle 249
```

```
Qy     583 CATTGGAGGATTAAAAGGCCAAATGCACCTGTGTGGATCGAATGGTATAGGAAGAGGCCC 642
           :::|||||||||:::||||||||||||          ||||||   |||   :::|||
Db     250 TyrTrpArgIleGlnArgProAsnAlaArgTrpTrpIleHisAlaTyrAsnArgArgThr 269

Qy     643 GACATGAATCCAGTAGTGTTGGAGCTTGCCATACTCGACTTAAATATTGTTCAAGCACAA 702
              :::||||||:::||||||||||||:::      ||||||  ||||||:::||||||||
Db     270 HisIleAsnProLeuValLeuGluLeuSerLysLeuAspPheAsnIleIleGlnAlaGln 289

Qy     703 TTTCAAGAAGAGCTCAAAGAATCCTTCAGGTGGTGGAGAAATACTGGGTTTGTTGAGAAG 762
           :::|||:::|||||||||:::          ||||||||||||||||||         ||||||
Db     290 TyrGlnGlnGluLeuLysGlnAspLeuArgTrpTrpArgAsnThrCysIleAlaGluLys 309

Qy     763 CTGCCCTTCGCAAGGGATAGACTGGTGGAATGCTACTTTTGGAATACTGGGATCATCGAG 822
           |||||||||||||||||||||||||||||     |||||||||:::||||||||||||:::
Db     310 LeuProPheAlaArgAspArgLeuValGluSerTyrPheTrpSerThrGlyIleIleGln 329

Qy     823 CCACGTCAGCATGCAAGTGCAAGGATAATGATGGGCAAAGTCAACGCTCTGATTACGGTG 882
           |||||||||||     :::||||||||||||||   |||      ||||||||||||
Db     330 ProArgGlnHisGluAsnAlaArgIleMetMetAlaLysAlaLeuAlaLeuIleThrThr 349

Qy     883 ATCGATGATATTTATGATGTCTATGGCACCTTAGAAGAACTCGAACAATTCACTGACCTC 942
           :::||||||:::|||||||||||||||||||||||||||||||||   |||    :::
Db     350 LeuAspAspValTyrAspValTyrGlyThrLeuGluGluLeuGluLeuPheIleGluAla 369

Qy     943 ATTCGAAGATGGGATATAAACTCAATCGACCAACTTCCCGATTACATGCAACTGTGCTTT 1002
           |||||||||||||:::|||:::||||||||||||||||:::|||||||||||||||||
Db     370 IleArgArgTrpGluIleSerSerIleAspGlnLeuProAsnTyrMetGlnLeuCysPhe 389

Qy    1003 CTTGCACTCAACAACTTCGTCGATGATACATCGTACGATGTTATGAAGGAGAAAGGCGTC 1062
           |||    :::|||||||||||||||||||||:::|||||||||||||||||||     :::
Db     390 LeuThrIleAsnAsnPheValAspAspThrAlaTyrAspValMetLysGluLysAspIle 409

Qy    1063 AACGTTATACCCTACCTGCGGCAATCGTGGGTTGATTTGGCGGATAAGTATATGGTAGAG 1122
           |||:::||||||||||||||:::||||||||||||||||||:::    |||:::||||||
Db     410 AsnIleIleProTyrLeuArgLysSerTrpValAspLeuAlaGluAlaTyrLeuValGlu 429

Qy    1123 GCACGGTGGTTCTACGGCGGGCACAAACCAAGTTTGGAAGAGTATTTGGAGAACTCATGG 1182
           |||:::||||||||||||||||||:::||||||||||||||||||    |||    |||
Db     430 AlaLysTrpPheTyrGlyGlyTyrLysProAsnLeuGluGluTyrLeuAsnAsnGlyTrp 449

Qy    1183 CAGTCGATAAGTGGGCCCTGTATGTTAACGCACATATTCTTCCGAGTAACAGATTCGTTC 1242
              |||:::|||||||||    :::|||   |||:::||||||   ||||||||||||
Db     450 IleSerValSerGlyProAlaIleLeuCysHisValPhePheGlyValThrAspSerIle 469

Qy    1243 ACAAAGGAGACCGTCGACAGTTTGTACAAATACCACGATTTAGTTCGTTGGTCATCCTTC 1302
           |||   |||||||||:::||||||:::|||||||||||||||:::|||   ||||||
Db     470 ThrMetGluThrValGluSerLeuPheLysTyrHisAspLeuIleArgCysSerSerThr 489

Qy    1303 GTTCTGCGGCTTGCTGATGATTTGGGAACCTCGGTGGAAGAGGTGAGCAGAGGGGATGTG 1362
           ::::::||||||||||||||||||      ||||||:::::||||||||||||||||||
Db     490 LeuValArgLeuAlaAspAspLeuAlaThrSerLeuAspGluValSerArgGlyAspVal 509

Qy    1363 CCGAAATCACTTCAGTGCTACATGAGTGACTACAATGCATCGGAGGCGGAGGCGCGGAAG 1422
           |||||||||:::||||||||||||||:::|||   ||||||||||||   |||||||||
Db     510 ProLysSerIleGlnCysTyrMetAsnAspAsnAsnAlaSerGluGluGluAlaArgLeu 529

Qy    1423 CACGTGAAATGGCTGATAGCGGAGGTGTGGAAGAAGATGAATGCGGAGAGGGTGTCGAAG 1482
           ||||||:::||||||||||||||||     ||||||:::||||||   |||   ||||||
Db     530 HisValArgTrpLeuIleAlaGluThrTrpLysGluMetAsnValGluMetValSerAla 549

Qy    1483 GATTCTCCATTCGGCAAAGATTTTATAGGATGTGCAGTTGATTTAGGAAGGATGGCGCAG 1542
           ||||||||||||    ||||||||||||||   ||||||    |||:::||||||||||||||||
Db     550 AspSerProPheCysLysAspPheIleAlaCysAlaAlaAspMetGlyArgMetAlaGln 569

Qy    1543 TTGATGTACCATAATGGAGATGGCACGGCACACAACACCCTATTATACATCAACAAATG 1602
              ||||||||||||||||||||||||||||   |||:::       ||||||||||||||||
Db     570 TyrMetTyrHisAsnGlyAspGlyHisGlyMetGlnAsnSerGlnIleHisGlnGlnMet 589

Qy    1603 ACCAGAACCTTATTCGAGCCCTTTGCA 1629
           |||    ||||||:::      |||
Db     590 ThrAspPheLeuPheGlnLysLeuAla 598
```

```
RESULT 13
O81191
ID   O81191      PRELIMINARY;      PRT;   591 AA.
AC   O81191;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   1,8-cineole synthase.
OS   Salvia officinalis (Sage).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC   lamiids; Lamiales; Lamiaceae; Nepetoideae; Mentheae; Salvia.
OX   NCBI_TaxID=38868;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98279000; PubMed=9614092;
RA   Wise M.L., Savage T.J., Katahira E., Croteau R.;
RT   "Monoterpene synthases from common sage (Salvia officinalis). cDNA
RT   isolation, characterization, and functional expression of (+)-sabinene
RT   synthase, 1,8-cineole synthase, and (+)-bornyl diphosphate synthase.";
RL   J. Biol. Chem. 273:14891-14899(1998).
DR   EMBL; AF051899; AAC26016.1; -.
DR   HSSP; Q40577; 5EAU.
DR   GO; GO:0016829; F:lyase activity; IEA.
DR   GO; GO:0008152; P:metabolism; IEA.
DR   InterPro; IPR005630; Terpene_synth_C.
DR   InterPro; IPR008949; Terpenoid_synth.
DR   InterPro; IPR008930; Terp_cyc_toroid.
DR   InterPro; IPR001906; Terp_synth-like.
DR   Pfam; PF01397; Terpene_synth; 1.
DR   Pfam; PF03936; Terpene_synth_C; 1.
SQ   SEQUENCE   591 AA;  69369 MW;  575F5C8A62C82D8B CRC64;

Alignment Scores:
Pred. No.:                5.04e-120     Length:         591
Score:                    1633.50       Matches:        305
Percent Similarity:       75.96%        Conservative:   109
Best Local Similarity:    55.96%        Mismatches:     116
Query Match:              56.17%        Indels:         15
DB:                       10            Gaps:           7

US-09-938-956-6 (1-1632) x O81191 (1-591)

Qy       4 AGACGATCCGGAAACTACAACCCTTCTCGTTGGGATGTCAACTTCATCCAATCGCTTCTC 63
           ||||||:::|||   |||   |||:::    ||||||    :::   ||||||
Db      57 ArgArgThrGlyGlyTyrGlnProThrLeuTrpAspPheSerThrIleGlnLeuPheAsp 76

Qy      64 AGTGACTATAAGGAGGACAAACACGTGATTAGGGCTTCTGAGCTGGTCACTTTGGTGAAG 123
           |||:::|||||||||:::||||||::::::||||||:::    ::::::      |||
Db      77 SerGluTyrLysGluGluLysHisLeuMetArgAlaAlaGlyMetIleAlaGlnValAsn 96

Qy     124 ATGGAACTGGAGAAAGAAACGGATCAAATTCGACAACTTGAGTTGATCGATGACTTGCAG 183
           |||   |||::::::|||   |||   |||::::::||||||||||||||||||||||:::
Db      97 MetLeuLeuGlnGluGluValAspSerIleGlnArgLeuGluLeuIleAspAspLeuArg 116

Qy     184 AGGATGGGGCTGTCCGATCATTTCCAAAATGAGTTCAAAGAAATCTTGTCCTCTATATAT 243
           |||:::|||   ||||||      |||      |||||||||:::|||      |||
Db     117 ArgLeuGlyIleSerCysHisPheAspArgGluIleValGluIleLeuAsnSerLysTyr 136

Qy     244 CTCGACCATCACTATTACAAGAACCCTTTTCCAAAAGAAGAAAGGGATCTCTACTCCACA 303
                       |||      |||            :::|||   |||||||||||||||
Db     137 ------------TyrThrAsnAsnGluIle------AspGluSerAspLeuTyrSerThr 150

Qy     304 TCTCTTGCATTTAGGCTCCTCAGAGAACATGGTTTTCAAGTCGCACAAGAGGTATTCGAT 363
           :::|||   |||:::|||||||||::::::   |||   |||:::||||||||||||
Db     151 AlaLeuArgPheLysLeuLeuArgGlnTyrAspPheSerValSerGlnGluValPheAsp 170

Qy     364 AGTTTCAAGAACGAGGAGGGT---GAGTTCAAAGAAAGCCTTAGCGACGACGACACCAGAGGA 420
              |||||||||::::::|||      :::||||||   ||||||   ||||||||||||||||
Db     171 CysPheLysAsnAspLysGlyThrAspPheLysProSerLeuValAspAspThrArgGly 190

Qy     421 TTGTTGCAACTGTATGAAGCTTCCTTTCTGTTGACGGAAGGCGAAACCACGCTCGAGTCA 480
           |||||||||||||||||||||||||||         :::||||||   ||||||
```

```
Db      191 LeuLeuGlnLeuTyrGluAlaSerPheLeuSerAlaGlnGlyGluGluThrLeuHisLeu 210

Qy      481 GCGAGGGAATTCGCCACCAAATTTTTGGAGGAAAAAGTGAACGAGGGTGGTGTTGATGGC 540
Db      211 AlaArgAspPheAlaThrLysPheLeuHisLysArgValLeu---------ValAspLys 227

Qy      541 GAC------CTTTTAACAAGAATCGCATATTCTTTGGACATCCCTCTTCATTGGAGGATT 594
Db      228 AspIleAsnLeuLeuSerSerIleGluArgAlaLeuGluLeuProThrHisTrpArgVal 247

Qy      595 AAAAGGCCAAATGCACCTGTGTGGATCGAATGGTATAGGAAGAGGCCCGACATGAATCCA 654
Db      248 GlnMetProAsnAlaArgSerPheIleAspAlaTyrLysArgArgProAspMetAsnPro 267

Qy      655 GTAGTGTTGGAGCTTGCCATACTCGACTTAAATATTGTTCAAGCACAATTTCAAGAAGAG 714
Db      268 ThrValLeuGluLeuAlaLysLeuAspPheAsnMetValGlnAlaGlnPheGlnGlnGlu 287

Qy      715 CTCAAAGAATCCTTCAGGTGGTGGAGAAATACTGGGTTTGTTGAGAAGCTGCCCTTCGCA 774
Db      288 LeuLysGluAlaSerArgTrpTrpAsnSerThrGlyLeuValHisGluLeuProPheVal 307

Qy      775 AGGGATAGACTGGTGGAATGCTACTTTTGGAATACTGGGATCATCGAGCCACGTCAGCAT 834
Db      308 ArgAspArgIleValGluCysTyrTyrTrpThrThrGlyValValGluArgArgGluHis 327

Qy      835 GCAAGTGCAAGGATAATGATGGGCAAAGTCAACGCTCTGATTACGGTGATCGATGATATT 894
Db      328 GlyTyrGluArgIleMetLeuThrLysIleAsnAlaLeuValThrThrIleAspAspVal 347

Qy      895 TATGATGTCTATGGCACCTTAGAAGAACTCGAACAATTCACTGACCTCATTCGAAGATGG 954
Db      348 PheAspIleTyrGlyThrLeuGluGluLeuGlnLeuPheThrThrAlaIleGlnArgTrp 367

Qy      955 GATATAAACTCAATCGACCAACTTCCCGATTACATGCAACTGTGCTTTCTTGCACTCAAC 1014
Db      368 AspIleGluSerMetLysGlnLeuProProTyrMetGlnIleCysTyrLeuAlaLeuPhe 387

Qy     1015 AACTTCGTCGATGATACATCGTACGATGTTATGAAGGAGAAAGGCGTCAACGTTATACCC 1074
Db      388 AsnPheValAsnGluMetAlaTyrAspThrLeuArgAspLysGlyPheAsnSerThrPro 407

Qy     1075 TACCTGCGGCAATCGTGGGTTGATTTGGCGGATAAGTATATGGTAGAGGCACGGTGGTTC 1134
Db      408 TyrLeuArgLysAlaTrpValAspLeuValGluSerTyrLeuIleGluAlaLysTrpTyr 427

Qy     1135 TACGGCGGGCACAAACCAAGTTTGGAAGAGTATTTGGAGAACTCATGGCAGTCGATAAGT 1194
Db      428 TyrMetGlyHisLysProSerLeuGluGluTyrMetLysAsnSerTrpIleSerIleGly 447

Qy     1195 GGGCCCTGTATGTTAACGCACATATTCTTCCGAGTAACAGATTCGTTCACAAAGGAGACC 1254
Db      448 GlyIleProIleLeuSerHisLeuPhePheArgLeuThrAspSerIleGluGluGluAsp 467

Qy     1255 GTCGACAGTTTGTACAAATACCACGATTTAGTTCGTTGGTCATCCTTCGTTCTGCGGCTT 1314
Db      468 AlaGluSerMetHisLysTyrHisAspIleValArgAlaSerCysThrIleLeuArgLeu 487

Qy     1315 GCTGATGATTTGGGAACCTCGGTGGAAGAGGTGAGCAGAGGGGATGTGCCGAAATCACTT 1374
Db      488 AlaAspAspMetGlyThrSerLeuAspGluValGluArgGlyAspValProLysSerVal 507

Qy     1375 CAGTGCTACATGAGTGACTACAATGCATCGGAGGCGGAGGCGCGGAAGCACGTGAAATGG 1434
Db      508 GlnCysTyrMetAsnGluLysAsnAlaSerGluGluGluAlaArgGluHisValArgSer 527

Qy     1435 CTGATAGCGGAGGTGTGGAAGAAGATGAATGCGGAGAGGGTGTCGAAGGATTCTCCATTC 1494
Db      528 LeuIleAspGlnThrTrpLysMetMetAsnLysGluMetMetThr------SerSerPhe 545

Qy     1495 GGCAAAGATTTTATAGGATGTGCAGTTGATTTAGGAAGGATGGCGCAGTTGATGTAC--- 1551
Db      546 SerLysTyrPheValGlnValSerAlaAsnLeuAlaArgMetAlaGlnTrpIleTyrGln 565

Qy     1552 CATAATGGAGATGGGCACGGCACACAACACCCTATTATACATCAACAAATGACCAGAACC 1611
Db      566 HisGluSerAspGlyPheGlyMetGlnHisSerLeuValAsnLysMetLeuArgGlyLeu 585

Qy     1612 TTATTCGAGCCCTTT 1626
Db      586 LeuPheAspArgTyr 590

RESULT 14
O81193
ID   O81193      PRELIMINARY;      PRT;   590 AA.
AC   O81193;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   (+)-sabinene synthase.
OS   Salvia officinalis (Sage).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC   lamiids; Lamiales; Lamiaceae; Nepetoideae; Mentheae; Salvia.
OX   NCBI_TaxID=38868;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98279000; PubMed=9614092;
RA   Wise M.L., Savage T.J., Katahira E., Croteau R.;
RT   "Monoterpene synthases from common sage (Salvia officinalis). cDNA
RT   isolation, characterization, and functional expression of (+)-sabinene
RT   synthase, 1,8-cineole synthase, and (+)-bornyl diphosphate synthase.";
RL   J. Biol. Chem. 273:14891-14899(1998).
DR   EMBL; AF051901; AAC26018.1; -.
DR   HSSP; Q40577; 5EAU.
DR   GO; GO:0016829; F:lyase activity; IEA.
DR   GO; GO:0008152; P:metabolism; IEA.
DR   InterPro; IPR005630; Terpene_synth_C.
DR   InterPro; IPR008949; Terpenoid_synth.
DR   InterPro; IPR008930; Terp_cyc_toroid.
DR   InterPro; IPR001906; Terp_synth-like.
DR   Pfam; PF01397; Terpene_synth; 1.
DR   Pfam; PF03936; Terpene_synth_C; 1.
SQ   SEQUENCE   590 AA;  68942 MW;  3E9CDFEDA2F0307D CRC64;

Alignment Scores:
Pred. No.:              1.01e-113     Length:         590
Score:                  1553.50       Matches:        291
Percent Similarity:     72.48%        Conservative:   104
Best Local Similarity:  53.39%        Mismatches:     141
Query Match:            53.42%        Indels:         9
DB:                     10            Gaps:           3

US-09-938-956-6 (1-1632) x O81193 (1-590)

Qy        1 ATGAGACGATCCGGAAACTACAACCCTTCTCGTTGGGATGTCAACTTCATCCAATCGCTT 60
Db       51 IleArgArgSerGlyAspTyrGlnProSerLeuTrpAspPheAsnTyrIleGlnSerLeu 70

Qy       61 CTCAGTGACTATAAGGAGGACAAACACGTGATTAGGGCTTCTGAGCTGGTCACTTTGGTG 120
Db       71 AsnThrProTyrLysGluGlnArgHisPheAsnArgGlnAlaGluLeuIleMetGlnVal 90

Qy      121 AAGATGGAACTGGAGAAAGAAACGGATCAAATTCGACAACTTGAGTTGATCGATGACTTG 180
Db       91 ArgMetLeuLeuLysValLysMetGluAlaIleGlnGlnLeuGluLeuIleAspAspLeu 110

Qy      181 CAGAGGATGGGGCTGTCCGATCATTTCCAAAATGAGTTCAAAGAAATCTTGTCCTCTATA 240
Db      111 GlnTyrLeuGlyLeuSerTyrPhePheGlnAspGluIleLysGlnIleLeuSerSerIle 130

Qy      241 TATCTCGACCATCACTATTACAAGAACCCTTTTCCAAAAGAAGAAAGGGATCTCTACTCC 300
Db      131 HisAsnGluProArgTyrPheHisAsn-----------------AsnAspLeuTyrPhe 144

Qy      301 ACATCTCTTGCATTTAGGCTCCTCAGAGAACATGGTTTTCAAGTCGCACAAGAGGTATTC 360
```

```
              |||:::|||  ||||||:::||||||:::|||||||||  |||:::::::::||||||
Db      145 ThrAlaLeuGlyPheArgIleLeuArgGlnHisGlyPheAsnValSerGluAspValPhe 164

Qy      361 GATAGTTTCAAGAACGAGGAG---GGTGAGTTCAAAGAAAGCCTTAGCGACGACACCAGA 417
            |||  ||||||   |||:::      :::|||        :::|||:::  |||||:::
Db      165 AspCysPheLysIleGluLysCysSerAspPheAsnAlaAsnLeuAlaGlnAspThrLys 184

Qy      418 GGATTGTTGCAACTGTATGAAGCTTCCTTTCTGTTGACGGAAGGCGAAACCACGCTCGAG 477
            |||:::|||||||||||||||||||||||||||  |||||||||  |||||||||
Db      185 GlyMetLeuGlnLeuTyrGluAlaSerPheLeuLeuArgGluGlyGluAspThrLeuGlu 204

Qy      478 TCAGCGAGGGAATTCGCCACCAAATTTTTGGAGGAAAAAGTGAACGAGGGTGGT------ 531
               ||||||   |||:::|||:::   |||   ||||||   :::|||||||||
Db      205 LeuAlaArgArgPheSerThrArgSerLeuArgGluLysPheAspGluGlyGlyAspGlu 224

Qy      532 GTTGATGGCGACCTTTTAACAAGAATCGCATATTCTTTGGACATCCCTCTTCATTGGAGG 591
            :::|||   ||||||   :::   |||   :::|||||||||:::|||||||||||||||
Db      225 IleAspGluAspLeuSerSerTrpIleArgHisSerLeuAspLeuProLeuHisTrpArg 244

Qy      592 ATTAAAAGGCCAAATGCACCTGTGTGGATCGAATGGTATAGGAAGAGGCCCGACATGAAT 651
            ::::::         |||      :::::::::   |||   :::|||||||||||||||
Db      245 ValGlnGlyLeuGluAlaArgTrpPheLeuAspAlaTyrAlaArgArgProAspMetAsn 264

Qy      652 CCAGTAGTGTTGGAGCTTGCCATACTCGACTTAAATATTGTTCAAGCACAATTTCAAGAA 711
            |||::::::   :::||||||   |||:::   |||||||||||||||   :::||||||
Db      265 ProLeuIlePheLysLeuAlaLysLeuAsnPheAsnIleValGlnAlaThrTyrGlnGlu 284

Qy      712 GAGCTCAAAGAATCCTTCAGGTGGTGGAGAAATACTGGGTTTGTTGAGAAGCTGCCCTTC 771
            |||||||||:::         |||||||||   ::::::      |||||||||||||||
Db      285 GluLeuLysAspIleSerArgTrpTrpAsnSerSerCysLeuAlaGluLysLeuProPhe 304

Qy      772 GCAAGGGATAGACTGGTGGAATGCTACTTTTGGAATACTGGGATCATCGAGCCACGTCAG 831
               |||||||||:::|||||||||:::||||||             ||||||   |||
Db      305 ValArgAspArgIleValGluCysPhePheTrpAlaIleAlaAlaPheGluProHisGln 324

Qy      832 CATGCAAGTGCAAGGATAATGATGGGCAAAGTCAACGCTCTGATTACGGTGATCGATGAT 891
            ::::::      |||   |||         :::         ||||||:::|||||||||
Db      325 TyrSerTyrGlnArgLysMetAlaAlaValIleIleThrPheIleThrIleIleAspAsp 344

Qy      892 ATTTATGATGTCTATGGCACCTTAGAAGAACTCGAACAATTCACTGACCTCATTCGAAGA 951
            :::||||||||||||||||||:::|||||||||||         ||||||:::|||||||||
Db      345 ValTyrAspValTyrGlyThrIleGluGluLeuGluLeuLeuThrAspMetIleArgArg 364

Qy      952 TGGGATATAAACTCAATCGACCAACTTCCCGATTACATGCAACTGTGCTTTCTTGCACTC 1011
            ||||||      ||||||   |||||||||   ||||||||||:::|||:::|||||||||
Db      365 TrpAspAsnLysSerIleSerGlnLeuProTyrTyrMetGlnValCysTyrLeuAlaLeu 384

Qy     1012 AACAACTTCGTCGATGATACATCGTACGATGTTATGAAGGAGAAAGGCGTCAACGTTATA 1071
               |||||||||   :::   :::||||||||:::::::|||:::::::   |||   |||
Db      385 TyrAsnPheValSerGluArgAlaTyrAspIleLeuLysAspGlnHisPheAsnSerIle 404

Qy     1072 CCCTACCTGCGGCAATCGTGGGTTGATTTGGCGGATAAGTATATGGTAGAGGCACGGTGG 1131
            |||||||||::::::|||||||||   |||   :::   |||:::   ||||||   |||
Db      405 ProTyrLeuGlnArgSerTrpValSerLeuValGluGlyTyrLeuLysGluAlaTyrTrp 424

Qy     1132 TTCTACGGCGGGCACAAACCAAGTTTGGAAGAGTATTTGGAGAACTCATGGCAGTCGATA 1191
            :::|||   |||:::|||||||||||||||||||||||||   |||:::      ||||||
Db      425 TyrTyrAsnGlyTyrLysProSerLeuGluGluTyrLeuAsnAsnAlaLysIleSerIle 444

Qy     1192 AGTGGGCCCTGTATGTTAACGCACATATTCTTCCGAGTAACAGATTCGTTCACAAAGGAG 1251
            |||   |||   :::::::::   :::::::|||   :::   :::|||         :::
Db      445 SerAlaProThrIleIleSerGlnLeuTyrPheThrLeuAlaAsnSerIleAspGluThr 464

Qy     1252 ACCGTCGACAGTTTGTACAAATACCACGATTTAGTTCGTTGGTCATCCTTCGTTCTGCGG 1311
               ::::::|||||||||:::||||||:::::::::         |||      :::||||||
Db      465 AlaIleGluSerLeuTyrGlnTyrHisAsnIleLeuTyrLeuSerGlyThrIleLeuArg 484

Qy     1312 CTTGCTGATGATTTGGGAACCTCGGTGGAAGAGGTGAGCAGAGGGGATGTGCCGAAATCA 1371
            ||||||||||||||||||||||||         |||:::   ||||||||||||||||||:::
Db      485 LeuAlaAspAspLeuGlyThrSerGlnHisGluLeuGluArgGlyAspValProLysAla 504

Qy     1372 CTTCAGTGCTACATGAGTGACTACAATGCATCGGAGGCGGAGGCGCGGAAGCACGTGAAA 1431
            :::|||||||||||||:::|||   ||||||||||||   ||||||   :::|||||||||
Db      505 IleGlnCysTyrMetAsnAspThrAsnAlaSerGluArgGluAlaValGluHisValLys 524

Qy     1432 TGGCTGATAGCGGAGGTGTGGAAGAAGATGAATGCGGAGAGGGTGTCGAAGGATTCTCCA 1491
            :::||||||   |||   ||||||:::||||||            :::   |||   |||
Db      525 PheLeuIleArgGluAlaTrpLysGluMetAsnThrValThrThrAlaSerAspCysPro 544

Qy     1492 TTCGGCAAAGATTTTATAGGATGTGCAGTTGATTTAGGAAGGATGGCGCAGTTGATGTAC 1551
            |||      |||   :::         |||   :::|||   |||   ||||||   :::|||
Db      545 PheThrAspAspLeuValAlaAlaAlaAlaAsnLeuAlaArgAlaAlaGlnPheIleTyr 564

Qy     1552 CATAATGGAGATGGGCACGGCACACAACACCCTATTATACATCAACAAATGACCAGAACC 1611
               :::|||||||||||||||   ||||||         |||||||||||||||
Db      565 LeuAspGlyAspGlyHisGlyValGlnHisSerGluIleHisGlnGlnMetGlyGlyLeu 584

Qy     1612 TTATTCGAGCCCTTT 1626
               ||||||:::|||:::
Db      585 LeuPheGlnProTyr 589
```

```
RESULT 15
Q8W1J9
ID   Q8W1J9      PRELIMINARY;      PRT;   607 AA.
AC   Q8W1J9;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Linalool synthase.
OS   Perilla frutescens var. crispa.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC   lamiids; Lamiales; Lamiaceae; Nepetoideae; Elsholtzieae; Perilla.
OX   NCBI_TaxID=179837;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. No. 79;
RA   Yamada N., Ito M., Honda G.;
RT   "Linalool synthase from Perilla frutescens, cloning and functional
RT   expression.";
RL   Submitted (NOV-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF444798; AAL38029.1; -.
DR   GO; GO:0016829; F:lyase activity; IEA.
DR   GO; GO:0008152; P:metabolism; IEA.
DR   InterPro; IPR005630; Terpene_synth_C.
DR   InterPro; IPR008949; Terpenoid_synth.
DR   InterPro; IPR008930; Terp_cyc_toroid.
DR   InterPro; IPR001906; Terp_synth-like.
DR   Pfam; PF01397; Terpene_synth; 1.
DR   Pfam; PF03936; Terpene_synth_C; 1.
SQ   SEQUENCE   607 AA;  71114 MW;  F3B309F5327A45D3 CRC64;

Alignment Scores:
Pred. No.:              1.44e-110      Length:         607
Score:                  1513.50        Matches:        292
Percent Similarity:     72.08%         Conservative:   103
Best Local Similarity:  53.28%         Mismatches:     144
Query Match:            52.05%         Indels:         9
DB:                     10             Gaps:           6

US-09-938-956-6 (1-1632) x Q8W1J9 (1-607)

Qy        4 AGACGATCCGGAAACTACAACCCTTCTCGTTGGGATGTCAACTTCATCCAATCGCTT--- 60
            ||||||||||||||||||   ||||||   ||||||   |||:::::::|||||||||
Db       63 ArgArgSerGlyAsnTyrGlnProSerLeuTrpAspPheAsnTyrLeuGlnSerLeuAsn 82

Qy       61 CTCAGTGACTATAAGGAGGACAAACACGTGATTAGGGCTTCTGAGCTGGTCACTTTGGTG 120
               :::   |||||||||   :::|||:::   |||   :::||||||:::      |||
Db       83 ThrThrGlnTyrLysGluValArgHisLeuLysArgGluAlaGluLeuIleGluGlnVal 102

Qy      121 AAGATGGAACTGGAGAAAGAAACGGATCAAATTCGACAACTTGAGTTGATCGATGACTTG 180
            ||||||   ||||||:::|||   :::   ::::::||||||||||||:::|||||||||
Db      103 LysMetLeuLeuGluGluGluMetGluAlaValGlnGlnLeuGluLeuValAspAspLeu 122

Qy      181 CAGAGGATGGGGCTGTCCGATCATTTCCAAAATGAGTTCAAAGAAATCTTGTCCTCTATA 240
            :::   :::|||||||||         |||:::::::::   |||:::||||||:::   |||
```

```
Db        123 LysAsnLeuGlyLeuSerTyrPhePheGluAspGlnIleLysGlnIleLeuThrPheIle 142

Qy        241 TATCTCGACCATCAC---TATTACAAGAACCCTTTT------CCAAAAGAAGAAAGGGAT 291
Db        143 TyrAsnGluHisLysCysPheHisSerAsnSerIleIleGluAlaGluGluIleArgAsp 162

Qy        292 CTCTACTCCACATCTCTTGCATTTAGGCTCCTCAGAGAACATGGTTTTCAAGTCGCACAA 351
Db        163 LeuTyrPheThrAlaLeuGlyPheArgLeuLeuArgGlnHisGlyPheGlnValSerGln 182

Qy        352 GAGGTATTCGATAGTTTCAAGAACGAGGAGGGT---GAGTTCAAAGAAAGCCTTAGCGAC 408
Db        183 GluValPheAspCysPheLysAsnGluGluGlySerAspPheLysAlaArgLeuGlyAsp 202

Qy        409 GACACCAGAGGATTGTTGCAACTGTATGAAGCTTCCTTTCTGTTGACGGAAGGCGAAACC 468
Db        203 AspThrLysGlyLeuLeuGlnLeuTyrGluAlaSerPheLeuLeuArgGluGlyGluAsp 222

Qy        469 ACGCTCGAGTCAGCGAGGGAATTCGCCACCAAATTTTTGGAGGAAAAAGTGAACGAGGGT 528
Db        223 ThrLeuGluLeuAlaArgGlnTyrAlaThrLysPheLeuGlnLysLysValAspHisGlu 242

Qy        529 GGTGTT---GATGGCGACCTTTTAACAAGAATCGCATATTCTTTGGACATCCCTCTTCAT 585
Db        243 LeuIleAspAspAsnAsnLeuLeuSerTrpIleLeuHisSerLeuGluIleProLeuHis 262

Qy        586 TGGAGGATTAAAAGGCCAAATGCACCTGTGTGGATCGAATGGTATAGGAAGAGGCCCGAC 645
Db        263 TrpArgIleGlnArgLeuGluAlaArgTrpPheLeuAspArgTyrAlaThrArgArgAsp 282

Qy        646 ATGAATCCAGTAGTGTTGGAGCTTGCCATACTCGACTTAAATATTGTTCAAGCACAATTT 705
Db        283 MetAsnGlnIleIleLeuGluLeuAlaLysLeuAspPheAsnIleIleGlnAlaThrGln 302

Qy        706 CAAGAAGAGCTCAAAGAATCCTTCAGGTGGTGGAGAAATACTGGGTTTGTTGAGAAGCTG 765
Db        303 GlnGluGluLeuLysAspLeuSerArgTrpTrpLysSerThrCysLeuAlaGluLysLeu 322

Qy        766 CCCTTCGCAAGGGATAGACTGGTGGAATGCTACTTTTGGAATACTGGGATCATCGAGCCA 825
Db        323 ProPheValArgAspArgLeuValGluSerTyrPheTrpAlaIleAlaLeuPheGluPro 342

Qy        826 CGTCAGCATGCAAGTGCAAGGATAATGATGGGCAAAGTCAACGCTCTGATTACGGTGATC 885
Db        343 HisGlnTyrGlyTyrHisArgLysValAlaAlaLysIleIleThrLeuIleThrSerLeu 362

Qy        886 GATGATATTTATGATGTCTATGGCACCTTAGAAGAACTCGAACAATTCACTGACCTCATT 945
Db        363 AspAspValTyrAspIleTyrGlyThrLeuAspGluLeuGlnLeuPheThrAspAlaIle 382

Qy        946 CGAAGATGGGATATAAACTCAATCGACCAACTTCCCGATTACATGCAACTGTGCTTTCTT 1005
Db        383 GlnArgTrpAspThrGluSerIleSerArgLeuProTyrTyrMetGlnLeuPheTyrMet 402

Qy       1006 GCACTCAACAACTTCGTCGATGATACATCGTACGATGTTATGAAGGAGAAAGGCGTCAAC 1065
Db        403 ValLeuTyrAsnPheValSerGluLeuAlaTyrAspGlyLeuLysGluLysGlyPheIle 422

Qy       1066 GTTATACCCTACCTGCGGCAATCGTGGGTTGATTTGGCGGATAAGTATATGGTAGAGGCA 1125
Db        423 ThrIleProTyrLeuGlnArgSerTrpAlaAspLeuValGluAlaTyrLeuLysGluAla 442

Qy       1126 CGGTGGTTCTACGGCGGGCACAAACCAAGTTTGGAAGAGTATTTGGAGAACTCATGGCAG 1185
Db        443 LysTrpPheTyrAsnGlyTyrValProSerMetGluGluTyrLeuAsnAsnAlaTyrIle 462

Qy       1186 TCGATAAGTGGGCCCTGTATGTTAACGCACATATTCTTCCGAGTAACAGATTCGTTCACA 1245
Db        463 SerIleGlyAlaThrProValIleSerGlnValPhePheThrLeuAlaThrSerIleAsp 482

Qy       1246 AAGGAGACCGTCGACAGTTTGTACAAATACCACGATTTAGTTCGTTGGTCATCCTTCGTT 1305
Db        483 LysProValIleAspSerLeuTyrGluTyrHisArgIleLeuArgLeuSerGlyMetLeu 502

Qy       1306 CTGCGGCTTGCTGATGATTTGGGAACCTCGGTGGAAGAGGTGAGCAGAGGGGATGTGCCG 1365
Db        503 ValArgLeuProAspAspLeuGlyThrSerProPheGluMetLysArgGlyAspValPro 522

Qy       1366 AAATCACTTCAGTGCTACATGAGTGACTACAATGCATCGGAGGCGGAGGCGCGGAAGCAC 1425
Db        523 LysAlaIleGlnLeuTyrMetLysGluArgAsnAlaThrGluIleGluAlaGlnGluHis 542

Qy       1426 GTGAAATGGCTGATAGCGGAGGTGTGGAAGAAGATGAATGCGGAGAGGGTGTCGAAGGAT 1485
Db        543 ValArgPheLeuIleArgGluAlaTrpLysGluMetAsnThrValThrThrAlaAlaAsp 562

Qy       1486 TCTCCATTCGGCAAAGATTTTATAGGATGTGCAGTTGATTTAGGAAGGATGGCGCAGTTG 1545
Db        563 CysProPheThrAspAspLeuValAlaAlaThrArgAsnLeuGlyArgAlaAlaGlnPhe 582

Qy       1546 ATGTACCATAATGGAGATGGGCACGGCACACAACACCCTATTATACATCAACAAATGACC 1605
Db        583 MetTyrLeuAspGlyAspGlyAsn---------HisSerGlnLeuHisGlnArgIleAla 599

Qy       1606 AGAACCTTATTCGAGCCCTTTGCA 1629
Db        600 CysLeuLeuPheGluProTyrAla 607
```

Search completed: February 25, 2004, 05:49:12
Job time : 120 secs

```
OM nucleic - protein search, using frame_plus_n2p model

Run on:          February 25, 2004, 05:36:43 ; Search time 21.5 Seconds
                                         (without alignments)
                                         7904.976 Million cell updates/sec

Title:           US-09-938-956-6
Perfect score:   2908
Sequence:        1 atgagacgatccggaaacta..........tattcgagccctttgcatga 1632

Scoring table:   BLOSUM62
                 Xgapop 10.0 , Xgapext  0.5
                 Ygapop 10.0 , Ygapext  0.5
                 Fgapop  6.0 , Fgapext  7.0
                 Delop   6.0 , Delext   7.0

Searched:        141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters:      283362

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US09938956/runat_24022004_144700_3053/app_query.fasta_1.1799
-DB=SwissProt_42 -QFMT=fastan -SUFFIX=rsp -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09938956_@CGN_1_1_21_@runat_24022004_144700_3053 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :       SwissProt_42:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	805.5	27.7	548	1	5EAS_TOBAC	Q40577 nicotiana t
	2	796	27.4	554	1	DCS2_GOSAR	Q39760 gossypium a
	3	781	26.9	554	1	DCS4_GOSAR	O49853 gossypium a
	4	776	26.7	554	1	DCS1_GOSAR	Q39761 gossypium a
	5	776	26.7	554	1	DCS1_GOSHI	P93665 gossypium h
	6	740	25.4	555	1	DCS3_GOSAR	Q43714 gossypium a
	7	714	24.6	628	1	TSD1_ABIGR	O24475 abies grand
	8	697	24.0	601	1	CASS_RICCO	P59287 ricinus com
	9	681.5	23.4	637	1	TSD3_ABIGR	O22340 abies grand
	10	672.5	23.1	627	1	TSD2_ABIGR	O24474 abies grand
	11	638	21.9	862	1	TASY_TAXBR	Q41594 taxus brevi
	12	635	21.8	862	1	TASY_TAXCH	Q9ft37 taxus chine
	13	634.5	21.8	862	1	TASY_TAXBA	Q93ya3 taxus bacca
c	14	138.5	4.8	1169	1	YK82_YEAST	P36170 saccharomyc
	15	130.5	4.5	2869	1	RBP1_PLAVB	Q00798 plasmodium
	16	127.5	4.4	943	1	SYL_METKA	Q8tvm4 methanopyru
	17	125.5	4.3	1030	1	Y075_MYCPN	P75556 mycoplasma
	18	124	4.3	1939	1	MYH6_MESAU	P13539 mesocricetu
	19	123.5	4.2	466	1	SYC_CLOPE	Q8xhq5 clostridium
	20	123	4.2	1713	1	LMA3_HUMAN	Q16787 homo sapien
	21	123	4.2	2672	1	GCN1_YEAST	P33892 saccharomyc
	22	122.5	4.2	4540	1	DYHC_PARTE	Q27171 paramecium
	23	120.5	4.1	1938	1	MYH6_RAT	P02563 rattus norv
	24	120.5	4.1	3680	1	DMD_CANFA	O97592 canis famil
	25	118	4.1	1875	1	MLP1_YEAST	Q02455 saccharomyc
c	26	117.5	4.0	725	1	AGA1_YEAST	P32323 saccharomyc
	27	117	4.0	3685	1	DMD_HUMAN	P11532 homo sapien
c	28	115	4.0	1140	1	YM96_YEAST	Q04893 saccharomyc
	29	115	4.0	1315	1	CHAO_DROME	P12024 drosophila
	30	114.5	3.9	795	1	LON_MYCGE	P47481 mycoplasma
	31	114.5	3.9	1938	1	MYH6_MOUSE	Q02566 mus musculu
	32	114.5	3.9	2555	1	FAFY_HUMAN	O00507 h probable
	33	114	3.9	847	1	SYA_HELPY	P56452 helicobacte
	34	114	3.9	1575	1	IQG2_HUMAN	Q13576 homo sapien
	35	113	3.9	1022	1	TPS3_YEAST	P38426 saccharomyc
	36	113	3.9	1587	1	SUR2_CAEEL	Q10669 caenorhabdi
	37	112	3.9	1163	1	YPT4_CAEEL	P41882 caenorhabdi
	38	112	3.9	8545	1	ANC1_CAEEL	Q9n4m4 caenorhabdi
	39	111.5	3.8	571	1	YD55_METJA	Q58750 methanococc
	40	111.5	3.8	1935	1	MYH7_HUMAN	P12883 homo sapien
	41	111.5	3.8	3744	1	TRA1_YEAST	P38811 saccharomyc
	42	111.5	3.8	6885	1	SNE2_HUMAN	Q8wxh0 homo sapien
	43	110.5	3.8	720	1	VP19_BPAPS	Q9t1s9 bacteriopha
	44	110.5	3.8	856	1	CNRB_MOUSE	P23440 mus musculu
	45	110.5	3.8	999	1	OXRP_RAT	Q63617 rattus norv

ALIGNMENTS

```
RESULT 1
5EAS_TOBAC
ID   5EAS_TOBAC     STANDARD;      PRT;   548 AA.
AC   Q40577;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Aristolchene synthase (EC 4.2.3.9) (5-epi-aristolochene synthase)
DE   (EAS).
OS   Nicotiana tabacum (Common tobacco).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC   lamiids; Solanales; Solanaceae; Nicotiana.
OX   NCBI_TaxID=4097;
RN   [1]
RP   SEQUENCE FROM N.A., AND SEQUENCE OF 56-73.
RC   STRAIN=cv. NK326;
RX   MEDLINE=93066390; PubMed=1438319;
RA   Facchini P.J., Chappell J.;
RT   "Gene family for an elicitor-induced sesquiterpene cyclase in
RT   tobacco.";
RL   Proc. Natl. Acad. Sci. U.S.A. 89:11088-11092(1992).
RN   [2]
RP   X-RAY CRYSTALLOGRAPHY (2.15 ANGSTROMS), AND REVISIONS.
RX   MEDLINE=97442533; PubMed=9295271;
RA   Starks C.M., Back K., Chappell J., Noel J.P.;
RT   "Structural basis for cyclic terpene biosynthesis by tobacco 5-epi-
RT   aristolochene synthase.";
RL   Science 277:1815-1820(1997).
CC   -!- FUNCTION: Catalyzes the cyclization of trans,trans-farnesyl
CC       diphosphate (FPP) to the bicyclic intermediate 5-epi-
CC       aristolochene, initial step in the conversion of FPP to the
CC       sesquiterpenoid antifungal phytoalexin capsidiol.
CC   -!- CATALYTIC ACTIVITY: Trans,trans-farnesyl diphosphate =
CC       aristolochene + diphosphate.
CC   -!- COFACTOR: Binds 3 magnesium ions per subunit.
CC   -!- SUBUNIT: Monomer.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -!- INDUCTION: By fungal elicitor.
CC   -!- SIMILARITY: Belongs to the terpene synthase family.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
```

```
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; L04680; AAA19216.1; -.
DR   PIR; T03714; T03714.
DR   PDB; 5EAS; 15-OCT-97.
DR   PDB; 5EAT; 12-NOV-97.
DR   PDB; 5EAU; 08-APR-98.
DR   InterPro; IPR008930; Terp_cyc_toroid.
DR   InterPro; IPR001906; Terp_synth-like.
DR   InterPro; IPR005630; Terpene_synth_C.
DR   InterPro; IPR008949; Terpenoid_synth.
DR   Pfam; PF01397; Terpene_synth; 1.
DR   Pfam; PF03936; Terpene_synth_C; 1.
KW   Lyase; Magnesium; 3D-structure.
FT   ACT_SITE     273    273       PROTON ACCEPTOR.
FT   ACT_SITE     444    444
FT   ACT_SITE     520    520
FT   ACT_SITE     525    525
FT   CONFLICT      42     42       Y -> YIY (IN REF. 1).
FT   CONFLICT      44     44       K -> Q (IN REF. 1).
FT   CONFLICT      55     55       N -> S (IN REF. 1).
FT   CONFLICT      62     62       M -> R (IN REF. 1).
FT   CONFLICT      73     73       T -> I (IN REF. 1).
FT   CONFLICT      89     89       D -> E (IN REF. 1).
FT   CONFLICT     388    388       T -> M (IN REF. 1).
FT   HELIX         26     29
FT   HELIX         36     57
FT   TURN          58     58
FT   TURN          60     61
FT   HELIX         64     76
FT   TURN          77     78
FT   HELIX         80     83
FT   HELIX         84     97
FT   HELIX        104    116
FT   TURN         117    118
FT   HELIX        123    129
FT   STRAND       130    130
FT   TURN         132    133
FT   STRAND       136    136
FT   HELIX        138    142
FT   HELIX        144    154
FT   TURN         155    156
FT   TURN         159    160
FT   HELIX        162    164
FT   TURN         165    166
FT   HELIX        167    178
FT   HELIX        179    181
FT   TURN         184    185
FT   HELIX        186    195
FT   TURN         199    200
FT   HELIX        203    213
FT   TURN         214    214
FT   HELIX        215    217
FT   TURN         219    220
FT   HELIX        223    254
FT   TURN         255    255
FT   HELIX        256    259
FT   TURN         261    262
FT   HELIX        267    277
FT   HELIX        281    283
FT   HELIX        284    306
FT   TURN         307    307
FT   HELIX        310    322
FT   TURN         323    323
FT   HELIX        326    330
FT   HELIX        333    354
FT   TURN         355    357
FT   HELIX        359    361
FT   HELIX        362    385
FT   TURN         386    386
FT   HELIX        391    398
FT   TURN         399    399
FT   HELIX        400    402
FT   HELIX        404    413
FT   TURN         414    414
FT   TURN         416    417
FT   HELIX        420    427
FT   TURN         428    428
FT   HELIX        431    454
FT   TURN         455    456
FT   TURN         458    459
FT   HELIX        461    469
FT   TURN         470    470
FT   HELIX        473    494
FT   TURN         495    495
FT   HELIX        503    505
FT   TURN         506    506
FT   HELIX        507    519
FT   TURN         520    520
FT   HELIX        535    542
FT   TURN         543    543
SQ   SEQUENCE   548 AA;  62973 MW;  9FE1C59CF1A68BF1 CRC64;

Alignment Scores:
Pred. No.:              4.35e-51      Length:         548
Score:                  805.50        Matches:        178
Percent Similarity:     52.92%        Conservative:   112
Best Local Similarity:  32.48%        Mismatches:     207
Query Match:            27.70%        Indels:         51
DB:                     1             Gaps:           10

US-09-938-956-6 (1-1632) x 5EAS_TOBAC (1-548)

Qy          1 ATGAGACGATCCGGAAACTACAACCCTTCTCGTTGGGATGTCAACTTCATCCAATCGCTT 60
                :::|||          :::::::::||||||    |||         |||:::
Db         14 ValArgProValAlaAspPheSerProSerLeuTrpGlyAspGlnPheLeuSerPheSer 33

Qy         61 CTCAGTGACTATAAGGAGGACAAACACGTGATTAGGGCTTCTGAGCTGGTCACTTTGGTG 120
              :::   :::          :::|||:::
Db         34 IleAspAsnGlnValAlaGluLysTyrAla------------------------------ 43

Qy        121 AAGATGGAACTGGAGAAAGAAACGGATCAAATTCGACAA--------------------- 159
                  |||:::|||          :::|||   |||
Db         44 ---LysGluIleGluAlaLeuLysGluGlnThrArgAsnMetLeuLeuAlaThrGlyMet 62

Qy        160 ---------------CTTGAGTTGATCGATGACTTGCAGAGGATGGGGCTGTCCGATCAT 204
                             |||   |||||||||   ::::::|||:::|||:::|||   |||
Db         63 LysLeuAlaAspThrLeuAsnLeuIleAspThrIleGluArgLeuGlyIleSerTyrHis 82

Qy        205 TTCCAAAATGAGTTCAAAGAAATCTTGTCCTCTATATATCTCGACCATCACTATTACAAG 264
              |||:::   |||      :::||||||      ||||||       :::
Db         83 PheGluLysGluIleAspAspIleLeuAspGlnIleTyrAsnGlnAsn------------ 98

Qy        265 AACCCTTTTCCAAAAGAAGAAAGGGATCTCTACTCCACATCTCTTGCATTTAGGCTCCTC 324
                                      ||||||   :::::::::|||   ||||||||||||
Db         99 ------------SerAsnCysAsnAspLeuCysThrSerAlaLeuGlnPheArgLeuLeu 114

Qy        325 AGAGAACATGGTTTTCAAGTCGCACAAGAGGTATTCGATAGTTTCAAGAACGAGGAGGGT 384
              |||:::|||||||||   ::::::   |||:::|||      |||::::::|||   |||
Db        115 ArgGlnHisGlyPheAsnIleSerProGluIlePheSerLysPheGlnAspGluAsnGly 134

Qy        385 GAGTTCAAAGAAAGCCTTAGCGACGACACCAGAGGATTGTTGCAACTGTATGAAGCTTCC 444
              :::|||||||||||||||||:::   |||      |||||||||   |||||||||||||||
Db        135 LysPheLysGluSerLeuAlaSerAspValLeuGlyLeuLeuAsnLeuTyrGluAlaSer 154

Qy        445 TTTCTGTTGACGGAAGGCGAAACCACGCTCGAGTCAGCGAGGGAATTCGCCACCAAATTT 504
                 :::   |||      :::      ||||||   |||      |||:::|||
Db        155 HisValArgThrHisAlaAspAspIleLeuGluAspAlaLeuAlaPheSerThrIleHis 174

Qy        505 TTGGAGGAAAAAGTGAACGAGGGTGGTGTTGATGGCGACCTTTTAACAAGAATCGCATAT 564
              ||||||                         :::      |||      ::::::   :::
```

```
Db     175 LeuGluSerAlaAlaProHis------LeuLysSerProLeuArgGluGlnValThrHis 192

Qy     565 TCTTTGGACATCCCTCTTCATTGGAGGATTAAAAGGCCAAATGCACCTGTGTGGATC--- 621
           :::|||:::      ||||||      :::  |||          :::|||
Db     193 AlaLeuGluGlnCysLeuHisLysGlyValProArgValGluThrArgPhePheIleSer 212

Qy     622 GAATGGTATAGGAAGAGGCCCGACATGAATCCAGTAGTGTTGGAGCTTGCCATACTCGAC 681
                 |||  |||          |||   |||:::|||   |||   ||||||
Db     213 SerIleTyrAspLysGluGlnSerLysAsnAsnValLeuLeuArgPheAlaLysLeuAsp 232

Qy     682 TTAAATATTGTTCAAGCACAATTTCAAGAAGAGCTCAAAGAATCCTTCAGGTGGTGGAGA 741
              |||::::::|||      ::::::||||||   :::      |||||||||:::
Db     233 PheAsnLeuLeuGlnMetLeuHisLysGlnGluLeuAlaGlnValSerArgTrpTrpLys 252

Qy     742 AATACTGGGTTTGTTGAGAAGCTGCCCTTCGCAAGGGATAGACTGGTGGAATGCTACTTT 801
           :::      ||||||      ||||||:::||||||||||||:::||||||||||||||
Db     253 AspLeuAspPheValThrThrLeuProTyrAlaArgAspArgValValGluCysTyrPhe 272

Qy     802 TGGAATACTGGGATCATCGAGCCACGTCAGCATGCAAGTGCAAGGATAATGATGGGCAAA 861
           |||      |||:::            |||::::::   ||||||:::|||:::   |||
Db     273 TrpAlaLeuGlyValTyrPheGluProGlnTyrSerGlnAlaArgValMetLeuValLys 292

Qy     862 GTCAACGCTCTGATTACGGTGATCGATGATATTTATGATGTCTATGGCACCTTAGAAGAA 921
              ::::::|||:::::::::||||||   :::|||   |||||||||::::::|||
Db     293 ThrIleSerMetIleSerIleValAspAspThrPheAspAlaTyrGlyThrValLysGlu 312

Qy     922 CTCGAACAATTCACTGACCTCATTCGAAGATGGGATATAAACTCAATCGACCAACTTCCC 981
           ||||||   :::||||||   |||:::||||||||||||||   ||||||:::||||||
Db     313 LeuGluAlaTyrThrAspAlaIleGlnArgTrpAspIleAsnGluIleAspArgLeuPro 332

Qy     982 GATTACATGCAACTGTGCTTTCTTGCACTCAACAACTTCGTCGATGATACATCGTACGAT 1041
           |||||||||::::::   :::   |||:::   :::         |||         :::
Db     333 AspTyrMetLysIleSerTyrLysAlaIleLeuAspLeuTyrLysAspTyrGluLysGlu 352

Qy    1042 GTTATGAAGGAGAAAGGCGTCAACGTTATACCCTACCTGCGGCAATCGTGGGTTGATTTG 1101
           :::         :::::::::   :::      :::         ::::::
Db     353 LeuSerSerAlaGlyArgSerHisIleValCysHisAlaIleGluArgMetLysGluVal 372

Qy    1102 GCGGATAAGTATATGGTAGAGGCACGGTGGTTCTACGGCGGGCACAAACCAAGTTTGGAA 1161
                 |||   ||||||:::   ||||||      |||:::   |||      :::
Db     373 ValArgAsnTyrAsnValGluSerThrTrpPheIleGluGlyTyrThrProProValSer 392

Qy    1162 GAGTATTTGGAGAACTCATGGCAGTCGATAAGTGGGCCCTGTATGTTAACGCACATATTC 1221
           |||||||||   |||:::      :::                   |||      :::
Db     393 GluTyrLeuSerAsnAlaLeuAlaThr---------------------ThrThrTyrTyr 405

Qy    1222 TTCCGAGTAACAGATTCGTTC------------------ACAAAGGAGACCGTCGACAGT 1263
           :::      |||   |||:::                  |||::::::      :::
Db     406 TyrLeuAlaThrThrSerTyrLeuGlyMetLysSerAlaThrGluGlnAspPheGluTrp 425

Qy    1264 TTGTACAAATACCACGATTTAGTTCGTTGGTCATCCTTCGTTCTGCGGCTTGCTGATGAT 1323
           |||   |||         ::::::      |||      :::   |||:::   ||||||
Db     426 LeuSerLysAsnProLysIleLeuGluAlaSerValIleIleCysArgValIleAspAsp 445

Qy    1324 TTGGGAACCTCGGTGGAAGAGGTGAGCAGAGGGGATGTGCCGAAATCACTTCAGTGCTAC 1383
              |||         |||   |||||||||   :::         ::::::|||
Db     446 ThrAlaThrTyrGluValGluLysSerArgGlyGlnIleAlaThrGlyIleGluCysCys 465

Qy    1384 ATGAGTGACTACAATGCATCGGAGGCGGAGGCGCGGAAGCACGTGAAATGGCTGATAGCG 1443
           |||   ||||||         |||      ||||||         :::   :::
Db     466 MetArgAspTyrGlyIleSerThrLysGluAlaMetAlaLysPheGlnAsnMetAlaGlu 485

Qy    1444 GAGGTGTGGAAGAAGATGAATGCGGAGAGGGTGTCGAAGGATTCTCCATTCGGCAAAGAT 1503
                 ||||||   :::|||   |||   :::   :::   :::|||         :::
Db     486 ThrAlaTrpLysAspIleAsn---GluGlyLeuLeuArgProThrProValSerThrGlu 504

Qy    1504 TTTATAGGATGTGCAGTTGATTTAGGAAGGATGGCGCAGTTGATGTAC---CATAATGGA 1560
           |||:::         ::::::|||   |||:::   ::::::   |||   ||||||
Db     505 PheLeuThrProIleLeuAsnLeuAlaArgIleValGluValThrTyrIleHisAsnLeu 524

Qy    1561 GATGGGCACGGCACACAACACCCT 1584
           ||||||:::         ||||||
Db     525 AspGlyTyr------ThrHisPro 530
```

```
RESULT 2
DCS2_GOSAR
ID   DCS2_GOSAR     STANDARD;      PRT;   554 AA.
AC   Q39760;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   (+)-delta-cadinene synthase isozyme XC14 (EC 4.2.3.13) (D-cadinene
DE   synthase).
OS   Gossypium arboreum (Tree cotton).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.
OX   NCBI_TaxID=29729;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Nanking;
RX   MEDLINE=96132653; PubMed=8554317;
RA   Chen X.-Y., Chen Y., Heinstein P., Davisson V.J.;
RT   "Cloning, expression, and characterization of (+)-delta-cadinene
RT   synthase: a catalyst for cotton phytoalexin biosynthesis.";
RL   Arch. Biochem. Biophys. 324:255-266(1995).
CC   -!- FUNCTION: Responsible for the cyclization of trans,trans-farnesyl
CC       diphosphate (FPP) to (+)-delta cadinene.
CC   -!- CATALYTIC ACTIVITY: 2-trans,6-trans-farnesyl diphosphate = (+)-
CC       delta-cadinene + diphosphate.
CC   -!- PATHWAY: Phytoalexins gossypol and lacinilene C biosynthesis;
CC       first (committed) step.
CC   -!- SIMILARITY: Belongs to the terpene synthase family.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute.  There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed.  Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; U23205; AAA93065.1; -.
DR   PIR; S68366; S68366.
DR   HSSP; Q40577; 5EAU.
DR   InterPro; IPR008930; Terp_cyc_toroid.
DR   InterPro; IPR001906; Terp_synth-like.
DR   InterPro; IPR005630; Terpene_synth_C.
DR   InterPro; IPR008949; Terpenoid_synth.
DR   Pfam; PF01397; Terpene_synth; 1.
DR   Pfam; PF03936; Terpene_synth_C; 1.
KW   Lyase; Multigene family.
FT   ACT_SITE    451    451       BY SIMILARITY.
FT   ACT_SITE    527    527       BY SIMILARITY.
FT   ACT_SITE    531    531       BY SIMILARITY.
SQ   SEQUENCE   554 AA;  64158 MW;  A88974665E0F6B2B CRC64;

Alignment Scores:
Pred. No.:              2.15e-50      Length:         554
Score:                  796.00        Matches:        185
Percent Similarity:     52.07%        Conservative:   104
Best Local Similarity:  33.33%        Mismatches:     232
Query Match:            27.37%        Indels:         34
DB:                     1             Gaps:           7

US-09-938-956-6 (1-1632) x DCS2_GOSAR (1-554)

Qy       1 ATGAGACGATCCGGAAACTACAACCCTTCTCGTTGG------------------------ 36
            ||||||       ::::::   ||||||   |||
Db      21 MetArgProLysAlaAspPheGlnProSerIleTrpGlyAspLeuPheLeuAsnCysPro 40

Qy      37 GATGTCAACTTCATCCAATCGCTTCTCAGTGACTATAAGGAGGACAAACACGTGATTAGG 96
            |||   |||                  :::::::::   |||      :::|||
Db      41 AspLysAsnIleAspAlaGluThrGluLysArgHisGlnGlnLeuLysGluGluValArg 60

Qy      97 GCTTCTGAGCTGGTCACTTTGGTGAAGATGGAACTGGAGAAAGAAACGGATCAAATTCGA 156
```

```
Db     61 -----------------------LysMetIleValAlaProMetAlaAsnSerThrGln 72
Qy    157 CAACTTGAGTTGATCGATGACTTGCAGAGGATGGGGCTGTCCGATCATTTCCAAAATGAG 216
Db     73 LysLeuAlaPheIleAspSerValGlnArgLeuGlyValSerTyrHisPheThrLysGlu 92
Qy    217 TTCAAAGAAATCTTGTCCTCTATATATCTCGACCATCACTATTACAAGAACCCTTTTCCA 276
Db     93 IleGluAspGluLeuGluAsnIleTyrHisAsnAsnAsn-------------------- 105
Qy    277 AAAGAAGAAAGGGATCTCTACTCCACATCTCTTGCATTTAGGCTCCTCAGAGAACATGGT 336
Db    106 AspAlaGluAsnAspLeuTyrThrThrSerLeuArgPheArgLeuLeuArgGluHisGly 125
Qy    337 TTTCAAGTCGCACAAGAGGTATTCGATAGTTTCAAGAACGAGGAGGGTGAGTTCAAAGAA 396
Db    126 PheAsnValSerCysAspValPheAsnLysPheLysAspGluGlnGlyAsnPheLysSer 145
Qy    397 AGCCTTAGCGACGACACCAGAGGATTGTTGCAACTGTATGAAGCTTCCTTTCTGTTGACG 456
Db    146 SerValThrSerAspValArgGlyLeuLeuGluLeuTyrGlnAlaSerTyrLeuArgVal 165
Qy    457 GAAGGCGAAACCACGCTCGAGTCAGCGAGGGAATTCGCCACCAAATTTTTGGAGGAAAAA 516
Db    166 HisGlyGluAspIleLeuAspGluAlaIleSerPheThrThrAsnHisLeuSerLeuAla 185
Qy    517 GTGAACGAGGGTGGTGTTGATGGCGACCTTTTAACAAGAATCGCATATTCTTTGGACATC 576
Db    186 Val------AlaSerLeuAspTyrProLeuSerGluGluValSerHisAlaLeuLysGln 203
Qy    577 CCTCTTCATTGGAGGATTAAAAGGCCAAATGCACCTGTGTGGATCGAATGGTATAGGAAG 636
Db    204 SerIleArgArgGlyLeuProArgValGluAlaArgHisTyrLeuSerValTyrGlnAsp 223
Qy    637 AGGCCCGACATGAATCCAGTAGTGTTGGAGCTTGCCATACTCGACTTAAATATTGTTCAA 696
Db    224 IleGluSerHisAsnLysValLeuLeuGluPheAlaLysIleAspPheAsnMetValGln 243
Qy    697 GCACAATTTCAAGAAGAGCTCAAAGAATCCTTCAGGTGGTGGAGAAATACTGGGTTTGTT 756
Db    244 LeuLeuHisArgLysGluLeuSerGluIleSerArgTrpTrpLysAspLeuAspPheGln 263
Qy    757 GAGAAGCTGCCCTTCGCAAGGGATAGACTGGTGGAATGCTACTTTTGGAATACTGGGATC 816
Db    264 ArgLysLeuProTyrAlaArgAspArgValValGluGlyTyrPheTrpIleSerGlyVal 283
Qy    817 ATCGAGCCACGTCAGCATGCAAGTGCAAGGATAATGATGGGCAAAGTCAACGCTCTGATT 876
Db    284 TyrPheGluProGlnTyrSerLeuGlyArgLysMetLeuThrLysValIleAlaMetAla 303
Qy    877 ACGGTGATCGATGATATTTATGATGTCTATGGCACCTTAGAAGAACTCGAACAATTCACT 936
Db    304 SerIleValAspAspThrTyrAspSerTyrAlaThrTyrGluGluLeuIleProTyrThr 323
Qy    937 GACCTCATTCGAAGATGGGATATAAACTCAATCGACCAACTTCCCGATTACATGCAACTG 996
Db    324 LysAlaIleGluArgTrpAspIleLysCysIleAspGluLeuProGluTyrMetLysPro 343
Qy    997 TGCTTTCTTGCACTCAACAACTTCGTCGATGATACATCGTACGATGTTATGAAGGAGAAA 1056
Db    344 SerTyrLysAlaLeuLeuAspValTyrGluGluMetGluGlnLeuValAlaLysHisGly 363
Qy   1057 GGCGTCAACGTTATACCCTACCTGCGGCAATCGTGGGTTGATTTGGCGGATAAGTATATG 1116
Db    364 ArgGlnTyrArgValGluTyrAlaLysAsnAlaMetIleArgLeuAlaGlnSerTyrLeu 383
Qy   1117 GTAGAGGCACGGTGGTTCTACGGCGGGCACAAACCAAGTTTGGAAGAGTATTTGGAGAAC 1176
Db    384 ValGluAlaArgTrpThrLeuGlnAsnTyrLysProSerPheGluGluPheLysAlaAsn 403
Qy   1177 TCATGGCAGTCGATAAGTGGGCCCTGTATGTTAACGCACATATTCTTCCGAGTAACAGAT 1236
Db    404 AlaLeuProThrCysGlyTyrAlaMetLeuAlaIleThrSerPheValGlyMetGlyAsp 423
Qy   1237 TCGTTCACAAAGGAGACCGTCGACAGTTTGTACAAATACCACGATTTAGTTCGTTGGTCA 1296
Db    424 IleValThrProGluThrPheLysTrpAlaAlaAsnAspProLysIleIleGlnAlaSer 443
Qy   1297 TCCTTCGTTCTGCGGCTTGCTGATGATTTGGGAACCTCGGTGGAAGAGGTGAGCAGAGGG 1356
Db    444 ThrIleIleCysArgPheMetAspAspValAlaGluHisLysPheLysHisArgArgGlu 463
Qy   1357 GATGTGCCGAAATCACTTCAGTGCTACATGAGTGACTACAATGCATCGGAGGCGGAGGCG 1416
Db    464 AspAspCysSerAlaIleGluCysTyrMetGluGluTyrGlyValThrAlaGlnGluAla 483
Qy   1417 ------------CGGAAGCACGTGAAATGGCTGATAGCGGAGGTGTGGAAGAAGATGAAT 1464
Db    484 TyrAspValPheAsnLysHisValGlu-----------SerAlaTrpLysAspValAsn 499
Qy   1465 GCGGAGAGGGTGTCGAAGGATTCTCCATTCGGCAAAGATTTTATAGGATGTGCAGTTGAT 1524
Db    500 LysGluPheLeu---LysProThrGluMetProThrGluValLeuAsnArgSerLeuAsn 518
Qy   1525 TTAGGAAGGATGGCGCAGTTGATGTACCATAATGGAGATGGGCACGGCACACAACACCCT 1584
Db    519 LeuAlaArgValMetAspValLeuTyrArgGluGlyAspGlyTyrThrTyrValGlyLys 538
Qy   1585 ATTATACATCAACAAATGACCAGAACCTTATTCGAGCCCTTTGCA 1629
Db    539 AlaAlaLysGlyGlyIleThrSerLeuLeuIleGluProValAla 553
```

```
RESULT 3
DCS4_GOSAR
ID   DCS4_GOSAR     STANDARD;      PRT;   554 AA.
AC   O49853;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   (+)-delta-cadinene synthase isozyme C2 (EC 4.2.3.13) (D-cadinene
DE   synthase).
GN   CAD1-C2.
OS   Gossypium arboreum (Tree cotton).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.
OX   NCBI_TaxID=29729;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Nanking;
RA   Meng Y., Jia J., Liu C., Liang W., Zhou X., Heinstein P., Chen X.-Y.;
RL   Submitted (FEB-1998) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Responsible for the cyclization of trans,trans-farnesyl
CC       diphosphate (FPP) to (+)-delta cadinene.
CC   -!- CATALYTIC ACTIVITY: 2-trans,6-trans-farnesyl diphosphate = (+)-
CC       delta-cadinene + diphosphate.
CC   -!- PATHWAY: Phytoalexins gossypol and lacinilene C biosynthesis;
CC       first (committed) step.
CC   -!- SIMILARITY: Belongs to the terpene synthase family.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; Y16432; CAA76223.1; -.
DR   HSSP; Q40577; 5EAU.
DR   InterPro; IPR008930; Terp_cyc_toroid.
DR   InterPro; IPR001906; Terp_synth-like.
DR   InterPro; IPR005630; Terpene_synth_C.
DR   InterPro; IPR008949; Terpenoid_synth.
DR   Pfam; PF01397; Terpene_synth; 1.
```

```
DR   Pfam; PF03936; Terpene_synth_C; 1.
KW   Lyase; Multigene family.
FT   ACT_SITE    451    451       BY SIMILARITY.
FT   ACT_SITE    527    527       BY SIMILARITY.
FT   ACT_SITE    531    531       BY SIMILARITY.
SQ   SEQUENCE   554 AA;  64117 MW;  35DDD66D3E838AAC CRC64;

Alignment Scores:
Pred. No.:               2.67e-49      Length:         554
Score:                   781.00        Matches:        184
Percent Similarity:      51.33%        Conservative:   105
Best Local Similarity:   32.68%        Mismatches:     224
Query Match:             26.86%        Indels:         50
DB:                      1             Gaps:           9

US-09-938-956-6 (1-1632) x DCS4_GOSAR (1-554)

Qy          1 ATGAGACGATCCGGAAACTACAACCCTTCTCGTTGG------------------------ 36
                :::|||        ::::::    ||||||  |||
Db         21 IleArgProLysAlaAspPheGlnProSerIleTrpGlyAspPhePheLeuAsnCysPro 40

Qy         37 GATGTCAACTTCATCCAATCGCTTCTCAGTGACTATAAGGAGGACAAACACGTGATTAGG 96
              |||   |||                             :::::::::   |||      :::|||
Db         41 AspLysAsnIleAspAlaGlyThrGluLysArgHisGlnGlnLeuLysGluGluValArg 60

Qy         97 GCTTCTGAGCTGGTCACTTTGGTGAAGATGGAACTGGAGAAAGAAACGGATCAAATTCGA 156
                                      ||||||   :::         :::       :::
Db         61 ------------------------LysMetIleValAlaProMetAlaAsnSerThrGln 72

Qy        157 CAACTTGAGTTGATCGATGACTTGCAGAGGATGGGGCTGTCCGATCATTTCCAAAATGAG 216
              :::|||      ||||||   :::||||||:::|||:::|||    ||||||      |||
Db         73 LysLeuAlaPheIleAspSerValGlnArgLeuGlyValSerTyrHisPheThrLysGlu 92

Qy        217 TTCAAAGAAATCTTGTCCTCTATATATCTCGACCATCACTATTACAAGAACCCTTTTCCA 276
              ::::::   |||    :::||||||   :::::::::
Db         93 IleGluAspGluLeuGluAsnIleTyrHisAsnAsnAsn--------------------- 105

Qy        277 AAAGAAGAAAGGGATCTCTACTCCACATCTCTTGCATTTAGGCTCCTCAGAGAACATGGT 336
                    |||   |||||||||:::|||||||||   |||||||||||||||||||||||||
Db        106 AspAlaGluAsnAspLeuTyrThrThrSerLeuArgPheArgLeuLeuArgGluHisGly 125

Qy        337 TTTCAAGTCGCACAAGAGGTATTCGATAGTTTCAAGAACGAGGAGGGTGAGTTCAAAGAA 396
              :::   |||:::   :::||||||:::   ||||||:::|||:::|||   ||||||
Db        126 TyrAsnValSerCysAspValPheAsnLysPheLysAspGluGlnGlyAsnPheLysSer 145

Qy        397 AGCCTTAGCGACGACACCAGAGGATTGTTGCAACTGTATGAAGCTTCCTTTCTGTTGACG 456
              |||::::::   |||   :::|||||||||:::||||||:::||||||:::|||
Db        146 SerValThrSerAspValGlnGlyLeuLeuGluLeuTyrGlnAlaSerTyrLeuArgVal 165

Qy        457 GAAGGCGAAACCACGCTCGAGTCAGCGAGGGAATTCGCCACCAAATTTTTGGAGGAAAAA 516
                 ||||||      |||:::   |||         |||   |||   |||
Db        166 HisGlyGluAspIleLeuAspGluAlaIleSerPheThrThrAsnHisLeuSerLeuAla 185

Qy        517 GTGAACGAGGGTGGTGTTGATGGCGACCTTTTAACAAGAATCGCATATTCTTTGGACATC 576
              |||:::         :::|||      |||         ::::::::::::|||
Db        186 ValSer------SerLeuAspHisProLeuSerGluGluValSerHisAlaLeuLysGln 203

Qy        577 CCTCTTCATTGGAGGATTAAAAGGCCAAATGCACCTGTGTGGATCGAATGGTATAGGAAG 636
                 :::      :::   |||         |||      ::::::      |||:::
Db        204 SerIleArgArgGlyLeuProArgValGluAlaArgHisTyrLeuSerValTyrGlnAsp 223

Qy        637 AGGCCCGACATGAATCCAGTAGTGTTGGAGCTTGCCATACTCGACTTAAATATTGTTCAA 696
                    |||      :::||||||   |||      :::|||   |||:::::::|||
Db        224 IleGluSerHisAsnLysAlaLeuLeuGluPheAlaLysIleAspPheAsnMetLeuGln 243

Qy        697 GCACAATTTCAAGAAGAGCTCAAAGAATCCTTCAGGTGGTGGAGAAATACTGGGTTTGTT 756
                    ::::::||||||   |||      |||||||||::::::         |||
Db        244 PheLeuHisArgLysGluLeuSerGluIleCysArgTrpTrpLysAspLeuAspPheGln 263

Qy        757 GAGAAGCTGCCCTTCGCAAGGGATAGACTGGTGGAATGCTACTTTTGGAATACTGGGATC 816
                 |||||||||:::||||||||||||:::||||||   |||||||||   :::|||:::
Db        264 ArgLysLeuProTyrAlaArgAspArgValValGluGlyTyrPheTrpIleSerGlyVal 283

Qy        817 ATCGAGCCACGTCAGCATGCAAGTGCAAGGATAATGATGGGCAAAGTCAACGCTCTGATT 876
                       |||::::::         |||   |||:::   ||||||   |||:::
Db        284 TyrPheGluProGlnTyrSerLeuGlyArgLysMetLeuThrLysValIleAlaMetAla 303

Qy        877 ACGGTGATCGATGATATTTATGATGTCTATGGCACCTTAGAAGAACTCGAACAATTCACT 936
              :::::::::||||||   ||||||   |||   |||   |||||||||      :::|||
Db        304 SerIleValAspAspThrTyrAspSerTyrAlaThrTyrGluGluLeuIleProTyrThr 323

Qy        937 GACCTCATTCGAAGATGGGATATAAACTCAATCGACCAACTTCCCGATTACATGCAACTG 996
              :::   |||   ||||||||||||      ||||||:::||||||:::|||||||:::
Db        324 AsnAlaIleGluArgTrpAspIleLysCysIleAspGluLeuProGluTyrMetLysPro 343

Qy        997 TGCTTTCTTGCACTCAACAACTTCGTCGATGATACATCGTACGATGTTATGAAGGAGAAA 1056
                 :::   ||||||                        ||||||   ||||||
Db        344 SerTyrLysAlaLeu------------------------LeuAspValTyrLysGluMet 355

Qy       1057 GGCGTCAACGTT------------------------ATACCCTACCTGCGGCAATCGTGG 1092
                    |||                              :::   |||   :::   :::
Db        356 GluGlnLeuValAlaGluHisGlyArgGlnTyrArgValGluTyrAlaLysAsnAlaMet 375

Qy       1093 GTTGATTTGGCGGATAAGTATATGGTAGAGGCACGGTGGTTCTACGGCGGGCACAAACCA 1152
              :::   ||||||      |||:::|||||||||||||||          :::||||||
Db        376 IleArgLeuAlaGlnSerTyrLeuValGluAlaArgTrpThrLeuGlnAsnTyrLysPro 395

Qy       1153 AGTTTGGAAGAGTATTTGGAGAACTCATGGCAGTCGATAAGTGGGCCCTGTATGTTAACG 1212
              |||   ||||||:::      |||:::      :::                  :::
Db        396 SerPheGluGluPheLysAlaAsnAlaLeuProThrCysGlyTyrAlaMetLeuAlaIle 415

Qy       1213 CACATATTCTTCCGAGTAACAGATTCGTTCACAAAGGAGACCGTCGACAGTTTGTACAAA 1272
                    |||      :::   |||      |||   ||||||
Db        416 ThrSerPheValGlyMetGlyAspIleValThrProGluThrPheLysTrpAlaAlaAsn 435

Qy       1273 TACCACGATTTAGTTCGTTGGTCATCCTTCGTTCTGCGGCTTGCTGATGATTTGGGAACC 1332
                 :::::::::   |||:::   :::      |||      ||||||:::
Db        436 AspProLysIleIleGlnAlaSerThrIleIleCysArgPheMetAspAspValAlaGlu 455

Qy       1333 TCGGTGGAAGAGGTGAGCAGAGGGGATGTGCCGAAATCACTTCAGTGCTACATGAGTGAC 1392
                    :::      |||   |||         :::::::::|||||||||      :::
Db        456 HisLysPheLysHisArgArgGluAspAspCysSerAlaIleGluCysTyrMetGluGlu 475

Qy       1393 TACAATGCATCGGAGGCGGAGGCG------------CGGAAGCACGTGAAATGGCTGATA 1440
              |||      |||      ||||||                  |||||||||:::
Db        476 TyrGlyValSerAlaGlnGluAlaTyrAspValPheAsnLysHisValGlu--------- 492

Qy       1441 GCGGAGGTGTGGAAGAAGATGAATGCGGAGAGGGTGTCGAAGGATTCTCCATTCGGCAAA 1500
                    ||||||   :::|||   |||                  |||   :::
Db        493 ---SerAlaTrpLysAspValAsnGlnGlu---PheGlnLysProThrGluMetProThr 510

Qy       1501 GATTTTATAGGATGTGCAGTTGATTTAGGAAGGATGGCGCAGTTGATGTACCATAATGGA 1560
              :::   :::      :::::::::|||   |||:::      ::::::|||      |||
Db        511 GluValLeuAsnArgSerLeuAsnLeuAlaArgValMetAspValLeuTyrArgGluGly 530

Qy       1561 GATGGGCACGGCACACAACACCCTATTATACATCAACAAATGACCAGAACCTTATTCGAG 1620
              ||||||:::                           :::|||         |||   |||
Db        531 AspGlyTyrThrTyrValGlyLysAlaAlaLysGlyGlyIleThrSerLeuLeuIleGlu 550

Qy       1621 CCCTTTGCA 1629
              |||   |||
Db        551 ProIleAla 553

RESULT 4
DCS1_GOSAR
ID   DCS1_GOSAR    STANDARD;     PRT;   554 AA.
AC   Q39761;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   (+)-delta-cadinene synthase isozyme XC1 (EC 4.2.3.13) (D-cadinene
DE   synthase).
OS   Gossypium arboreum (Tree cotton).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.
```

```
OX   NCBI_TaxID=29729;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Nanking;
RX   MEDLINE=96132653; PubMed=8554317;
RA   Chen X.-Y., Chen Y., Heinstein P., Davisson V.J.;
RT   "Cloning, expression, and characterization of (+)-delta-cadinene
RT   synthase: a catalyst for cotton phytoalexin biosynthesis.";
RL   Arch. Biochem. Biophys. 324:255-266(1995).
CC   -!- FUNCTION: Responsible for the cyclization of trans,trans-farnesyl
CC       diphosphate (FPP) to (+)-delta cadinene.
CC   -!- CATALYTIC ACTIVITY: 2-trans,6-trans-farnesyl diphosphate = (+)-
CC       delta-cadinene + diphosphate.
CC   -!- PATHWAY: Phytoalexins gossypol and lacinilene C biosynthesis;
CC       first (committed) step.
CC   -!- SIMILARITY: Belongs to the terpene synthase family.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; U23206; AAA93064.1; -.
DR   PIR; S68365; S68365.
DR   HSSP; Q40577; 5EAU.
DR   InterPro; IPR008930; Terp_cyc_toroid.
DR   InterPro; IPR001906; Terp_synth-like.
DR   InterPro; IPR005630; Terpene_synth_C.
DR   InterPro; IPR008949; Terpenoid_synth.
DR   Pfam; PF01397; Terpene_synth; 1.
DR   Pfam; PF03936; Terpene_synth_C; 1.
KW   Lyase; Multigene family.
FT   ACT_SITE    451    451       BY SIMILARITY.
FT   ACT_SITE    527    527       BY SIMILARITY.
FT   ACT_SITE    531    531       BY SIMILARITY.
SQ   SEQUENCE   554 AA;  64137 MW;  59D6922DEDF9DCAF CRC64;

Alignment Scores:
Pred. No.:              6.19e-49      Length:        554
Score:                  776.00        Matches:       178
Percent Similarity:     52.07%        Conservative:  111
Best Local Similarity:  32.07%        Mismatches:    232
Query Match:            26.69%        Indels:        34
DB:                     1             Gaps:          7

US-09-938-956-6 (1-1632) x DCS1_GOSAR (1-554)

Qy          1 ATGAGACGATCCGGAAACTACAACCCTTCTCGTTGG------------------------ 36
                ||||||          ::::::    ||||||    |||
Db         21 MetArgProLysAlaAspPheGlnProSerIleTrpGlyAspLeuPheLeuAsnCysPro 40

Qy         37 GATGTCAACTTCATCCAATCGCTTCTCAGTGACTATAAGGAGGACAAACACGTGATTAGG 96
                |||   |||                        :::::::::   |||    :::|||
Db         41 AspLysAsnIleAspAlaGluThrGluLysArgHisGlnGlnLeuLysGluGluValArg 60

Qy         97 GCTTCTGAGCTGGTCACTTTGGTGAAGATGGAACTGGAGAAAGAAACGGATCAAATTCGA 156
                                        ||||||   :::          :::      :::
Db         61 ------------------------LysMetIleValAlaProMetAlaAsnSerThrGln 72

Qy        157 CAACTTGAGTTGATCGATGACTTGCAGAGGATGGGGCTGTCCGATCATTTCCAAAATGAG 216
              :::|||       ||||||   :::||||||:::|||:::|||   ||||||      |||
Db         73 LysLeuAlaPheIleAspSerValGlnArgLeuGlyValSerTyrHisPheThrLysGlu 92

Qy        217 TTCAAAGAAATCTTGTCCTCTATATATCTCGACCATCACTATTACAAGAACCCTTTTCCA 276
                 ::::::    |||   :::||||||   :::::::::
Db         93 IleGluAspGluLeuGluAsnIleTyrHisAsnAsnAsn--------------------- 105

Qy        277 AAAGAAGAAAGGGATCTCTACTCCACATCTCTTGCATTTAGGCTCCTCAGAGAACATGGT 336
                    |||   |||||||||:::||||||:::      |||||||||||||||||||||
Db        106 AspAlaGluAsnAspLeuTyrThrThrSerIleArgPheArgLeuLeuArgGluHisGly 125

Qy        337 TTTCAAGTCGCACAAGAGGTATTCGATAGTTTCAAGAACGAGGAGGGTGAGTTCAAAGAA 396
              :::   |||:::   :::||||||:::      ||||||:::|||:::|||   ||||||
Db        126 TyrAsnValSerCysAspValPheAsnLysPheLysAspGluGlnGlyAsnPheLysSer 145

Qy        397 AGCCTTAGCGACGACACCAGAGGATTGTTGCAACTGTATGAAGCTTCCTTTCTGTTGACG 456
              |||::::::   |||   ||||||||||||:::|||||||:::||||||:::|||
Db        146 SerValThrSerAspValArgGlyLeuLeuGluLeuTyrGlnAlaSerTyrLeuArgVal 165

Qy        457 GAAGGCGAAACCACGCTCGAGTCAGCGAGGGAATTCGCCACCAAATTTTTGGAGGAAAAA 516
                 ||||||      |||:::   |||   |||:::   |||      |||
Db        166 HisGlyGluAspIleLeuAspGluAlaIleSerPheThrThrHisHisLeuSerLeuAla 185

Qy        517 GTGAACGAGGGTGGTGTTGATGGCGACCTTTTAACAAGAATCGCATATTCTTTGGACATC 576
              |||                   :::|||   |||          ::::::::::::|||
Db        186 Val------AlaSerLeuAspHisProLeuSerGluGluValSerHisAlaLeuLysGln 203

Qy        577 CCTCTTCATTGGAGGATTAAAAGGCCAAATGCACCTGTGTGGATCGAATGGTATAGGAAG 636
                 :::         :::   |||       |||      ::::::      |||:::
Db        204 SerIleArgArgGlyLeuProArgValGluAlaArgHisTyrLeuSerValTyrGlnAsp 223

Qy        637 AGGCCCGACATGAATCCAGTAGTGTTGGAGCTTGCCATACTCGACTTAAATATTGTTCAA 696
                       |||      :::||||||   |||   :::|||   |||::::::|||
Db        224 IleGluSerHisAsnLysAlaLeuLeuGluPheAlaLysIleAspPheAsnMetLeuGln 243

Qy        697 GCACAATTTCAAGAAGAGCTCAAAGAATCCTTCAGGTGGTGGAGAAATACTGGGTTTGTT 756
                    ::::::||||||   |||      ||||||||||::::::      |||
Db        244 PheLeuHisArgLysGluLeuSerGluIleCysArgTrpTrpLysAspLeuAspPheGln 263

Qy        757 GAGAAGCTGCCCTTCGCAAGGGATAGACTGGTGGAATGCTACTTTTGGAATACTGGGATC 816
                 |||||||||:::||||||||||||:::||||||   |||||||||   :::|||:::
Db        264 ArgLysLeuProTyrAlaArgAspArgValValGluGlyTyrPheTrpIleSerGlyVal 283

Qy        817 ATCGAGCCACGTCAGCATGCAAGTGCAAGGATAATGATGGGCAAAGTCAACGCTCTGATT 876
                       |||::::::      |||   |||:::   ||||||   |||:::
Db        284 TyrPheGluProGlnTyrSerLeuGlyArgLysMetLeuThrLysValIleAlaMetAla 303

Qy        877 ACGGTGATCGATGATATTTATGATGTCTATGGCACCTTAGAAGAACTCGAACAATTCACT 936
              :::::::::||||||   ||||||   |||   |||   |||||||||      :::|||
Db        304 SerIleValAspAspThrTyrAspSerTyrAlaThrTyrGluGluLeuIleProTyrThr 323

Qy        937 GACCTCATTCGAAGATGGGATATAAACTCAATCGACCAACTTCCCGATTACATGCAACTG 996
              :::   |||   ||||||||||||         ||||||::::::|||:::|||||||:::
Db        324 AsnAlaIleGluArgTrpAspIleLysCysIleAspGluIleProGluTyrMetLysPro 343

Qy        997 TGCTTTCTTGCACTCAACAACTTCGTCGATGATACATCGTACGATGTTATGAAGGAGAAA 1056
                 :::   ||||||   :::   ::::::                  |||   :::
Db        344 SerTyrLysAlaLeuLeuAspValTyrGluGluMetValGlnLeuValAlaGluHisGly 363

Qy       1057 GGCGTCAACGTTATACCCTACCTGCGGCAATCGTGGGTTGATTTGGCGGATAAGTATATG 1116
                       :::   |||   :::   :::   :::   ||||||      |||:::
Db        364 ArgGlnTyrArgValGluTyrAlaLysAsnAlaMetIleArgLeuAlaGlnSerTyrLeu 383

Qy       1117 GTAGAGGCACGGTGGTTCTACGGCGGGCACAAACCAAGTTTGGAAGAGTATTTGGAGAAC 1176
              |||||||||:::|||               :::|||||||||   ||||||:::   |||
Db        384 ValGluAlaLysTrpThrLeuGlnAsnTyrLysProSerPheGluGluPheLysAlaAsn 403

Qy       1177 TCATGGCAGTCGATAAGTGGGCCCTGTATGTTAACGCACATATTCTTCCGAGTAACAGAT 1236
              :::   :::               :::               |||   :::   |||
Db        404 AlaLeuProThrCysGlyTyrAlaMetLeuAlaIleThrSerPheValGlyMetGlyAsp 423

Qy       1237 TCGTTCACAAAGGAGACCGTCGACAGTTTGTACAAATACCACGATTTAGTTCGTTGGTCA 1296
                    |||   ||||||                     :::::::::   |||
Db        424 IleValThrProGluThrPheLysTrpAlaAlaSerAspProLysIleIleGlnAlaSer 443

Qy       1297 TCCTTCGTTCTGCGGCTTGCTGATGATTTGGGAACCTCGGTGGAAGAGGTGAGCAGAGGG 1356
              :::   :::   |||      ||||||:::               :::         |||
Db        444 ThrIleIleCysArgPheMetAspAspValAlaGluHisLysPheLysHisArgArgGlu 463

Qy       1357 GATGTGCCGAAATCACTTCAGTGCTACATGAGTGACTACAATGCATCGGAGGCGGAGGCG 1416
              |||         :::::::::|||||||||   :::|||      :::      ||||||
Db        464 AspAspCysSerAlaIleGluCysTyrMetGluGluTyrGlyValThrAlaGlnGluAla 483
```

```
Qy       1417 ------------CGGAAGCACGTGAAATGGCTGATAGCGGAGGTGTGGAAGAAGATGAAT 1464
                          |||||||||:::                 ||||||   :::|||
Db        484 TyrAspValPheAsnLysHisValGlu------------SerAlaTrpLysAspLeuAsn 499

Qy       1465 GCGGAGAGGGTGTCGAAGGATTCTCCATTCGGCAAAGATTTTATAGGATGTGCAGTTGAT 1524
                 |||   :::  |||   :::            :::  :::    :::::::::
Db        500 GlnGluPheLeu---LysProThrGluMetProThrGluValLeuAsnArgSerLeuAsn 518

Qy       1525 TTAGGAAGGATGGCGCAGTTGATGTACCATAATGGAGATGGGCACGGCACACAACACCCT 1584
               |||  |||:::       ::::::|||        |||||||||:::
Db        519 LeuAlaArgValMetAspValLeuTyrArgGluGlyAspGlyTyrThrTyrValGlyLys 538

Qy       1585 ATTATACATCAACAAATGACCAGAACCTTATTCGAGCCCTTTGCA 1629
                          :::|||      |||   ||||||   |||
Db        539 AlaAlaLysGlyGlyIleThrSerLeuLeuIleGluProIleAla 553
```

```
RESULT 5
DCS1_GOSHI
ID   DCS1_GOSHI     STANDARD;      PRT;   554 AA.
AC   P93665;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   (+)-delta-cadinene synthase (EC 4.2.3.13) (D-cadinene synthase).
GN   CDN1.
OS   Gossypium hirsutum (Upland cotton).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.
OX   NCBI_TaxID=3635;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Westburn M;
RA   Davis E.M., Chen Y.-S., Essenberg M., Pierce M.L.;
RT   "cDNA sequence of a (+)-delta-cadinene synthase gene induced in
RT   Gossypium hirsutum L. by bacterial infection.";
RL   (In) Plant Gene Register PGR98-040.
RN   [2]
RP   PARTIAL SEQUENCE, AND CHARACTERIZATION.
RX   MEDLINE=96351891; PubMed=8728715;
RA   Davis E.M., Tsuji J., Davis G.D., Pierce M.L., Essenberg M.;
RT   "Purification of (+)-delta-cadinene synthase, a sesquiterpene cyclase
RT   from bacteria-inoculated cotton foliar tissue.";
RL   Phytochemistry 41:1047-1055(1996).
CC   -!- FUNCTION: Responsible for the cyclization of trans,trans-farnesyl
CC       diphosphate (FPP) to (+)-delta cadinene.
CC   -!- CATALYTIC ACTIVITY: 2-trans,6-trans-farnesyl diphosphate = (+)-
CC       delta-cadinene + diphosphate.
CC   -!- PATHWAY: Phytoalexins gossypol and lacinilene C biosynthesis;
CC       first (committed) step.
CC   -!- INDUCTION: By bacterial infection.
CC   -!- SIMILARITY: Belongs to the terpene synthase family.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; U88318; AAC12784.1; -.
DR   HSSP; Q40577; 5EAU.
DR   InterPro; IPR008930; Terp_cyc_toroid.
DR   InterPro; IPR001906; Terp_synth-like.
DR   InterPro; IPR005630; Terpene_synth_C.
DR   InterPro; IPR008949; Terpenoid_synth.
DR   Pfam; PF01397; Terpene_synth; 1.
DR   Pfam; PF03936; Terpene_synth_C; 1.
KW   Lyase.
FT   ACT_SITE    451    451       BY SIMILARITY.
FT   ACT_SITE    527    527       BY SIMILARITY.
FT   ACT_SITE    531    531       BY SIMILARITY.
SQ   SEQUENCE   554 AA;  64019 MW;  8BCC78AD8CA5B816 CRC64;

  Alignment Scores:
  Pred. No.:              6.19e-49      Length:         554
  Score:                  776.00        Matches:        181
  Percent Similarity:     51.71%        Conservative:   106
  Best Local Similarity:  32.61%        Mismatches:     234
  Query Match:            26.69%        Indels:         34
  DB:                     1             Gaps:           7

US-09-938-956-6 (1-1632) x DCS1_GOSHI (1-554)
```

```
Qy          1 ATGAGACGATCCGGAAACTACAACCCTTCTCGTTGG------------------------ 36
              ||||||         ::::::   ||||||   |||
Db         21 MetArgProLysAlaAspPheGlnProSerIleTrpGlyAspPhePheLeuAsnCysPro 40

Qy         37 GATGTCAACTTCATCCAATCGCTTCTCAGTGACTATAAGGAGGACAAACACGTGATTAGG 96
              |||   |||                           :::::::::   |||      :::|||
Db         41 AspLysAsnIleAspAlaGluThrGlnLysArgHisGlnGlnLeuLysGluGluValArg 60

Qy         97 GCTTCTGAGCTGGTCACTTTGGTGAAGATGGAACTGGAGAAAGAAACGGATCAAATTCGA 156
                                    ||||||   :::             :::
Db         61 -----------------------LysMetIleValAlaProMetAlaAsnSerThrLeu 72

Qy        157 CAACTTGAGTTGATCGATGACTTGCAGAGGATGGGGCTGTCCGATCATTTCCAAAATGAG 216
              :::|||      ||||||   :::|||   :::|||:::|||   ||||||      |||
Db         73 LysLeuAlaPheIleAspSerValGlnGlyLeuGlyValSerTyrHisPheThrLysGlu 92

Qy        217 TTCAAAGAAATCTTGTCCTCTATATATCTCGACCATCACTATTACAAGAACCCTTTTCCA 276
                 ::::::   |||   :::||||||   :::::::::
Db         93 IleGluAspGluLeuGluAsnIleTyrHisAsnAsnAsn--------------------- 105

Qy        277 AAAGAAGAAAGGGATCTCTACTCCACATCTCTTGCATTTAGGCTCCTCAGAGAACATGGT 336
                       |||   |||||||||:::|||||||||   |||||||||||||||||||||
Db        106 AspAlaGluAsnAspLeuTyrThrThrSerLeuArgPheArgLeuLeuArgGluHisGly 125

Qy        337 TTTCAAGTCGCACAAGAGGTATTCGATAGTTTCAAGAACGAGGAGGGTGAGTTCAAAGAA 396
              |||   |||:::   :::||||||:::   ||||||:::|||:::|||   ||||||
Db        126 PheHisValSerCysAspValPheAsnLysPheLysAspGluGlnGlyAsnPheLysSer 145

Qy        397 AGCCTTAGCGACGACACCAGAGGATTGTTGCAACTGTATGAAGCTTCCTTTCTGTTGACG 456
              |||::::::   |||   ||||||||||||:::||||||:::||||||:::|||
Db        146 SerValThrSerAspValArgGlyLeuLeuGluLeuTyrGlnAlaSerTyrLeuArgVal 165

Qy        457 GAAGGCGAAACCACGCTCGAGTCAGCGAGGGAATTCGCCACCAAATTTTTGGAGGAAAAA 516
              ||||||      ||||:::      |||         |||      :::      |||
Db        166 HisGlyGluAspIleLeuAspGluAlaIleSerPheThrSerAsnHisLeuSerLeuAla 185

Qy        517 GTGAACGAGGGTGGTGTTGATGGCGACCTTTTAACAAGAATCGCATATTCTTTGGACATC 576
              |||            :::|||      |||            ::::::::::::|||
Db        186 Val------AlaSerLeuAspHisProLeuSerGluGluValSerHisAlaLeuLysGln 203

Qy        577 CCTCTTCATTGGAGGATTAAAAGGCCAAATGCACCTGTGTGGATCGAATGGTATAGGAAG 636
                 :::      :::   |||      |||      ::::::      |||:::
Db        204 SerIleArgArgGlyLeuProArgValGluAlaArgHisTyrLeuSerValTyrGlnAsp 223

Qy        637 AGGCCCGACATGAATCCAGTAGTGTTGGAGCTTGCCATACTCGACTTAAATATTGTTCAA 696
                       |||   |||:::||||||   |||      :::|||   |||:::||||||
Db        224 IleGluSerHisAsnLysValLeuLeuGluPheAlaLysIleAspPheAsnMetValGln 243

Qy        697 GCACAATTTCAAGAAGAGCTCAAAGAATCCTTCAGGTGGTGGAGAAATACTGGGTTTGTT 756
                 ::::::||||||   |||      |||||||||::::::            |||
Db        244 LeuLeuHisArgLysGluLeuSerGluIleSerArgTrpTrpLysAspLeuAspPheGln 263

Qy        757 GAGAAGCTGCCCTTCGCAAGGGATAGACTGGTGGAATGCTACTTTTGGAATACTGGGATC 816
                 |||||||||:::|||||||||||||:::||||||   |||||||||   :::|||:::
Db        264 ArgLysLeuProTyrAlaArgAspArgValValGluGlyTyrPheTrpIleSerGlyVal 283

Qy        817 ATCGAGCCACGTCAGCATGCAAGTGCAAGGATAATGATGGGCAAAGTCAACGCTCTGATT 876
                       |||::::::      |||   |||:::   ||||||   |||:::
Db        284 TyrPheGluProGlnTyrSerLeuGlyArgLysMetLeuThrLysValIleAlaMetAla 303

Qy        877 ACGGTGATCGATGATATTTATGATGTCTATGGCACCTTAGAAGAACTCGAACAATTCACT 936
```

```
              :::::::::||||||  ||||||  |||  |||  |||||||||     :::|||
Db        304 SerIleValAspAspThrTyrAspSerTyrAlaThrTyrGluGluLeuIleProTyrThr 323

Qy        937 GACCTCATTCGAAGATGGGATATAAACTCAATCGACCAACTTCCCGATTACATGCAACTG 996
              :::  |||  |||||||||||      ||||||:::||||||:::||||||:::
Db        324 AsnAlaIleGluArgTrpAspIleLysCysIleAspGluLeuProGluTyrMetLysPro 343

Qy        997 TGCTTTCTTGCACTCAACAACTTCGTCGATGATACATCGTACGATGTTATGAAGGAGAAA 1056
              :::  ||||||  :::      ::::::                  |||  :::
Db        344 SerTyrLysAlaLeuLeuAspValTyrGluGluMetGluGlnLeuValAlaGluHisGly 363

Qy       1057 GGCGTCAACGTTATACCCTACCTGCGGCAATCGTGGGTTGATTTGGCGGATAAGTATATG 1116
                    :::  |||  :::  :::  :::  ||||||     |||:::
Db        364 ArgGlnTyrArgValGluTyrAlaLysAsnAlaMetIleArgLeuAlaGlnSerTyrLeu 383

Qy       1117 GTAGAGGCACGGTGGTTCTACGGCGGGCACAAACCAAGTTTGGAAGAGTATTTGGAGAAC 1176
              |||||||||||||||            :::|||||||||  ||||||:::      |||
Db        384 ValGluAlaArgTrpThrLeuGlnAsnTyrLysProSerPheGluGluPheLysAlaAsn 403

Qy       1177 TCATGGCAGTCGATAAGTGGGCCCTGTATGTTAACGCACATATTCTTCCGAGTAACAGAT 1236
              :::     :::          :::              |||  :::  |||
Db        404 AlaLeuProThrCysGlyTyrAlaMetLeuAlaIleThrSerPheValGlyMetGlyAsp 423

Qy       1237 TCGTTCACAAAGGAGACCGTCGACAGTTTGTACAAATACCACGATTTAGTTCGTTGGTCA 1296
                    |||  ||||||                  :::::::::      |||
Db        424 IleValThrProGluThrPheLysTrpAlaAlaAsnAspProLysIleIleGlnAlaSer 443

Qy       1297 TCCTTCGTTCTGCGGCTTGCTGATGATTTGGGAACCTCGGTGGAAGAGGTGAGCAGAGGG 1356
              :::  :::  |||     ||||||:::             :::    |||
Db        444 ThrIleIleCysArgPheMetAspAspValThrGluHisLysPheLysHisArgArgGlu 463

Qy       1357 GATGTGCCGAAATCACTTCAGTGCTACATGAGTGACTACAATGCATCGGAGGCGGAGGCG 1416
              |||       :::::::::|||||||||  :::|||     :::       ||||||
Db        464 AspAspCysSerAlaIleGluCysTyrMetGluGluTyrGlyValThrAlaGlnGluAla 483

Qy       1417 ------------CGGAAGCACGTGAAATGGCTGATAGCGGAGGTGTGGAAGAAGATGAAT 1464
                          ||||||||||:::                  ||||||  :::|||
Db        484 TyrAspValPheAsnLysHisValGlu------------SerAlaTrpLysAspValAsn 499

Qy       1465 GCGGAGAGGGTGTCGAAGGATTCTCCATTCGGCAAAGATTTTATAGGATGTGCAGTTGAT 1524
                 :::        |||  :::            :::  :::     :::::::::
Db        500 ---GlnGlyPheLeuLysProThrGluMetProThrGluValLeuAsnArgSerLeuAsn 518

Qy       1525 TTAGGAAGGATGGCGCAGTTGATGTACCATAATGGAGATGGGCACGGCACACAACACCCT 1584
              |||  |||:::      ::::::|||     |||||||||:::
Db        519 LeuAlaArgValMetAspValLeuTyrArgGluGlyAspGlyTyrThrTyrValGlyLys 538

Qy       1585 ATTATACATCAACAAATGACCAGAACCTTATTCGAGCCCTTTGCA 1629
                             :::|||     ||||||  |||
Db        539 AlaAlaLysGlyGlyIleThrSerLeuLeuIleGluProIleAla 553

RESULT 6
DCS3_GOSAR
ID   DCS3_GOSAR   .   STANDARD;      PRT;   555 AA.
AC   Q43714;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   (+)-delta-cadinene synthase isozyme A (EC 4.2.3.13) (D-cadinene
DE   synthase).
GN   CAD1-A.
OS   Gossypium arboreum (Tree cotton).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.
OX   NCBI_TaxID=29729;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Nanking;
RX   MEDLINE=97060798; PubMed=8904844;
RA   Chen X.-Y., Wang M., Chen Y., Davisson V.J., Heinstein P.;
RT   "Cloning and heterologous expression of a second (+)-delta-cadinene
RT   synthase from Gossypium arboreum.";
RL   J. Nat. Prod. 59:944-951(1996).
CC   -!- FUNCTION: Responsible for the cyclization of trans,trans-farnesyl
CC       diphosphate (FPP) to (+)-delta cadinene.
CC   -!- CATALYTIC ACTIVITY: 2-trans,6-trans-farnesyl diphosphate = (+)-
CC       delta-cadinene + diphosphate.
CC   -!- PATHWAY: Phytoalexins gossypol and lacinilene C biosynthesis;
CC       first (committed) step.
CC   -!- SIMILARITY: Belongs to the terpene synthase family.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; X96429; CAA65289.1; -.
DR   EMBL; U27535; AAB41259.1; -.
DR   HSSP; Q40577; 5EAT.
DR   InterPro; IPR008930; Terp_cyc_toroid.
DR   InterPro; IPR001906; Terp_synth-like.
DR   InterPro; IPR005630; Terpene_synth_C.
DR   InterPro; IPR008949; Terpenoid_synth.
DR   Pfam; PF01397; Terpene_synth; 1.
DR   Pfam; PF03936; Terpene_synth_C; 1.
KW   Lyase; Multigene family.
FT   ACT_SITE    452    452       BY SIMILARITY.
FT   ACT_SITE    528    528       BY SIMILARITY.
FT   ACT_SITE    532    532       BY SIMILARITY.
SQ   SEQUENCE   555 AA;  64103 MW;  7060C4F9D99412F9 CRC64;

Alignment Scores:
Pred. No.:               2.62e-46      Length:        555
Score:                   740.00        Matches:       174
Percent Similarity:      51.91%        Conservative:  111
Best Local Similarity:   31.69%        Mismatches:    238
Query Match:             25.45%        Indels:        26
DB:                      1             Gaps:          7

US-09-938-956-6 (1-1632) x DCS3_GOSAR (1-555)

Qy          4 AGACGATCCGGAAACTACAACCCTTCTCGTTGGGATGTCAACTTCATCCAATCGCTTCTC 63
              |||       :::::::::|||     |||        ||||||
Db         21 ArgProLysAlaAspPheHisProGlyIleTrpGlyAspMetPheIleIleCysProAsp 40

Qy         64 AGTGAC------------------------TATAAGGAGGACAAACACGTGATTAGGGCT 99
              :::|||                        |||:::|||  |||     :::|||
Db         41 ThrAspIleAspAlaAlaThrGluLeuGlnTyrGluGluLeuLysAlaGlnValArg--- 59

Qy        100 TCTGAGCTGGTCACTTTGGTGAAGATGGAACTGGAGAAAGAAACGGATCAAATTCGACAA 159
                                   ||||||  :::|||        |||       ::::::
Db         60 ---------------------LysMetIleMetGluProValAspAspSerAsnGlnLys 72

Qy        160 CTTGAGTTGATCGATGACTTGCAGAGGATGGGGCTGTCCGATCATTTCCAAAATGAGTTC 219
              |||     ||||||  :::||||||:::|||:::|||  |||||||:::   |||
Db         73 LeuProPheIleAspAlaValGlnArgLeuGlyValSerTyrHisPheGluLysGluIle 92

Qy        220 AAAGAAATCTTGTCCTCTATATATCTCGACCATCACTATTACAAGAACCCTTTTCCAAAA 279
              ::::::  |||  :::||||||  |||  :::
Db         93 GluAspGluLeuGluAsnIleTyrArgAspThrAsnAsn---------------AsnAsp 107

Qy        280 GAAGAAAGGGATCTCTACTCCACATCTCTTGCATTTAGGCTCCTCAGAGAACATGGTTTT 339
              :::  |||||||||:::|||:::|||  ||||||||||||||||||||||||
Db        108 AlaAspThrAspLeuTyrThrThrAlaLeuArgPheArgLeuLeuArgGluHisGlyPhe 127

Qy        340 CAAGTCGCACAAGAGGTATTCGATAGTTTCAAGAACGAGGAGGGTGAGTTCAAAGAAAGC 399
              ::::::  :::  |||:::  ||||||:::|||  |||  ||||||  |||
Db        128 AspIleSerCysAspAlaPheAsnLysPheLysAspGluAlaGlyAsnPheLysAlaSer 147

Qy        400 CTTAGCGACGACACCAGAGGATTGTTGCAACTGTATGAAGCTTCCTTTCTGTTGACGGAA 459
              |||:::  |||  :::|||||||||:::|||||||||||||||:::::
Db        148 LeuThrSerAspValGlnGlyLeuLeuGluLeuTyrGluAlaSerTyrMetArgValHis 167
```

```
Qy        460 GGCGAAACCACGCTCGAGTCAGCGAGGGAATTCGCCACCAAATTTTTGGAGGAAAAAGTG 519
              ||||||      |||:::      |||      |||   |||        |||       :::
Db        168 GlyGluAspIleLeuAspGluAlaIleSerPheThrThrAlaGlnLeuThrLeuAlaLeu 187

Qy        520 AACGAGGGTGGTGTTGATGGCGACCTTTTAACAAGAATCGCATATTCTTTGGACATCCCT 579
                                :::             |||     ::::::   ::::::|||
Db        188 ProThr------LeuHisHisProLeuSerGluGlnValGlyHisAlaLeuLysGlnSer 205

Qy        580 CTTCATTGGAGGATTAAAAGGCCAAATGCACCTGTGTGGATCGAATGGTATAGGAAGAGG 639
              :::      :::   |||      |||      :::|||      |||:::
Db        206 IleArgArgGlyLeuProArgValGluAlaArgAsnPheIleSerIleTyrGlnAspLeu 225

Qy        640 CCCGACATGAATCCAGTAGTGTTGGAGCTTGCCATACTCGACTTAAATATTGTTCAAGCA 699
                 |||      :::|||:::   |||      :::|||   |||::::::|||
Db        226 GluSerHisAsnLysSerLeuLeuGlnPheAlaLysIleAspPheAsnLeuLeuGlnLeu 245

Qy        700 CAATTTCAAGAAGAGCTCAAAGAATCCTTCAGGTGGTGGAGAAATACTGGGTTTGTTGAG 759
                 ::::::||||||   |||      |||||||||::::::         |||
Db        246 LeuHisArgLysGluLeuSerGluIleCysArgTrpTrpLysAspLeuAspPheThrArg 265

Qy        760 AAGCTGCCCTTCGCAAGGGATAGACTGGTGGAATGCTACTTTTGGAATACTGGGATCATC 819
              ||||||||||||||||||||||||:::||||||   |||||||||      |||:::
Db        266 LysLeuProPheAlaArgAspArgValValGluGlyTyrPheTrpIleMetGlyValTyr 285

Qy        820 GAGCCACGTCAGCATGCAAGTGCAAGGATAATGATGGGCAAAGTCAACGCTCTGATTACG 879
                 |||::::::      |||   |||:::   ||||||   |||:::   :::
Db        286 PheGluProGlnTyrSerLeuGlyArgLysMetLeuThrLysValIleAlaMetAlaSer 305

Qy        880 GTGATCGATGATATTTATGATGTCTATGGCACCTTAGAAGAACTCGAACAATTCACTGAC 939
              ::::::||||||      ||||||   |||   |||   :::||||||      :::|||:::
Db        306 IleValAspAspThrTyrAspSerTyrAlaThrTyrAspGluLeuIleProTyrThrAsn 325

Qy        940 CTCATTCGAAGATGGGATATAAACTCAATCGACCAACTTCCCGATTACATGCAACTGTGC 999
                 |||   ||||||||||||         ::::::|||||||||:::||||||::::::
Db        326 AlaIleGluArgTrpAspIleLysCysMetAsnGlnLeuProAsnTyrMetLysIleSer 345

Qy       1000 TTTCTTGCACTCAACAACTTCGTCGATGATACATCGTACGATGTTATGAAGGAGAAAGGC 1059
              :::   ||||||   |||      ::::::            :::      :::
Db        346 TyrLysAlaLeuLeuAsnValTyrGluGluMetGluGlnLeuLeuAlaAsnGlnGlyArg 365

Qy       1060 GTCAACGTTATACCCTACCTGCGGCAATCGTGGGTTGATTTGGCGGATAAGTATATGGTA 1119
                 :::   |||   :::::::::   :::   |||         |||::::::
Db        366 GlnTyrArgValGluTyrAlaLysLysAlaMetIleArgLeuValGlnAlaTyrLeuLeu 385

Qy       1120 GAGGCACGGTGGTTCTACGGCGGGCACAAACCAAGTTTGGAAGAGTATTTGGAGAACTCA 1179
              ||||||:::|||   :::      :::||||||:::   ||||||:::   :::|||:::
Db        386 GluAlaLysTrpThrHisGlnAsnTyrLysProThrPheGluGluPheArgAspAsnAla 405

Qy       1180 TGGCAGTCGATAAGTGGGCCCTGTATGTTAACG---CACATATTCTTCCGAGTAACAGAT 1236
                 :::   ||||||   ||||||         |||         :::   :::
Db        406 LeuProThr---SerGlyTyrAlaMetLeuAlaIleThrAlaPheValGlyMetGlyGlu 424

Qy       1237 TCGTTCACAAAGGAGACCGTCGACAGTTTGTACAAATACCACGATTTAGTTCGTTGGTCA 1296
                 |||   ||||||                  :::::::::         |||
Db        425 ValIleThrProGluThrPheLysTrpAlaAlaSerAspProLysIleIleLysAlaSer 444

Qy       1297 TCCTTCGTTCTGCGGCTTGCTGATGATTTGGGAACCTCGGTGGAAGAGGTGAGCAGAGGG 1356
              :::   :::   |||      ||||||:::                        |||
Db        445 ThrIleIleCysArgPheMetAspAspIleAlaGluHisLysPheAsnHisArgArgGlu 464

Qy       1357 GATGTGCCGAAATCACTTCAGTGCTACATGAGTGACTACAATGCATCGGAGGCGGAGGCG 1416
              |||         :::::::::|||||||||      |||      :::      ||||||
Db        465 AspAspCysSerAlaIleGluCysTyrMetLysGlnTyrGlyValThrAlaGlnGluAla 484

Qy       1417 CGGAAGCACGTGAAATGGCTGATAGCGGAGGTGTGGAAGAAGATGAATGCGGAGAGGGTG 1476
                             |||         ||||||   :::|||   |||
Db        485 TyrAsnGluPheAsnLysHisIleGluSerSerTrpLysAspValAsn---GluGluPhe 503

Qy       1477 TCGAAGGATTCTCCATTCGGCAAAGATTTTATAGGATGTGCAGTTGATTTAGGAAGGATG 1536
                 |||   :::                  :::      :::::::::|||   |||:::
Db        504 LeuLysProThrGluMetProThrProValLeuCysArgSerLeuAsnLeuAlaArgVal 523

Qy       1537 GCGCAGTTGATGTACCATAATGGAGATGGGCACGGCACACAACACCCTATTATACATCAA 1596
                       ::::::|||         |||||||||:::
Db        524 MetAspValLeuTyrArgGluGlyAspGlyTyrThrHisValGlyLysAlaAlaLysGly 543

Qy       1597 CAAATGACCAGAACCTTATTCGAGCCC 1623
                 :::|||      |||   :::|||
Db        544 GlyIleThrSerLeuLeuIleAspPro 552
```

```
RESULT 7
TSD1_ABIGR
ID   TSD1_ABIGR     STANDARD;      PRT;   628 AA.
AC   O24475;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Pinene synthase, chloroplast precursor (EC 4.2.3.14) (Beta-geraniolene
DE   synthase) ((-)-(1S,5S)-pinene synthase).
GN   AG3.18.
OS   Abies grandis (Grand fir).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Abies.
OX   NCBI_TaxID=46611;
RN   [1]
RP   SEQUENCE FROM N.A., AND CHARACTERIZATION.
RX   MEDLINE=97413772; PubMed=9268308;
RA   Bohlmann J., Steele C.L., Croteau R.;
RT   "Monoterpene synthases from grand fir (Abies grandis). cDNA isolation,
RT   characterization, and functional expression of myrcene synthase, (-)-
RT   (4S)-limonene synthase, and (-)-(1S,5S)-pinene synthase.";
RL   J. Biol. Chem. 272:21784-21792(1997).
CC   -!- FUNCTION: Involved in defensive oleoresin formation in conifers in
CC       response to insect attack or other injury. Involved in monoterpene
CC       (C10) olefins biosynthesis. A mixture of alpha-and beta-pinene is
CC       produced by this enzyme.
CC   -!- CATALYTIC ACTIVITY: Geranyl diphosphate = pinene + diphosphate.
CC   -!- COFACTOR: Manganese and potassium.
CC   -!- PATHWAY: Oleoresinosis.
CC   -!- SUBCELLULAR LOCATION: Chloroplast.
CC   -!- INDUCTION: By wounding.
CC   -!- SIMILARITY: Belongs to the terpene synthase family.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; U87909; AAB71085.1; -.
DR   HSSP; Q40577; 5EAT.
DR   InterPro; IPR008930; Terp_cyc_toroid.
DR   InterPro; IPR001906; Terp_synth-like.
DR   InterPro; IPR005630; Terpene_synth_C.
DR   InterPro; IPR008949; Terpenoid_synth.
DR   Pfam; PF01397; Terpene_synth; 1.
DR   Pfam; PF03936; Terpene_synth_C; 1.
KW   Lyase; Manganese; Transit peptide; Chloroplast.
FT   TRANSIT       1      ?       CHLOROPLAST (POTENTIAL).
FT   CHAIN         ?    628       PINENE SYNTHASE.
FT   ACT_SITE    524    524       BY SIMILARITY.
FT   ACT_SITE    601    601       BY SIMILARITY.
FT   ACT_SITE    605    605       BY SIMILARITY.
SQ   SEQUENCE   628 AA;  71505 MW;  23DBB788F3C8072C CRC64;

Alignment Scores:
Pred. No.:                2.12e-44      Length:          628
Score:                    714.00        Matches:         191
Percent Similarity:       52.19%        Conservative:    107
Best Local Similarity:    33.45%        Mismatches:      231
Query Match:              24.55%        Indels:          42
DB:                       1             Gaps:            14
```

```
US-09-938-956-6 (1-1632) x TSD1_ABIGR (1-628)

Qy          4 AGACGATCCGGAAACTACAACCCTTCTCGTTGGGATGTCAACTTCATCCAATCGCTTCTC 63
                ||||||   |||:::::::::   :::    |||||    :::   |||||||||||
Db         64 ArgArgMetGlyAspPheHisSerAsnLeuTrpAspAspAspValIleGlnSerLeuPro 83

Qy         64 AGTGACTATAAGGAGGACAAACACGTGATTAGGGCTTCTGAGCTGGTCACTTTGGTGAAG 123
              :::    |||:::|||      ::::::    ||||||   :::|||:::        ||||||
Db         84 ThrAlaTyrGluGluLysSerTyrLeuGluArgAlaGluLysLeuIleGlyGluValLys 103

Qy        124 ---------------ATGGAACTGGAGAAAGAA-----------------ACGGATCAA 150
                             |||   ||||||                        |||
Db        104 AsnMetPheAsnSerMetSerLeuGluAspGlyGluLeuMetSerProLeuAsnAspLeu 123

Qy        151 ATTCGACAACTTGAGTTGATCGATGACTTGCAGAGGATGGGGCTGTCCGATCATTTCCAA 210
              |||::::::|||   ::::::|||   |||:::|||:::|||:::        ||||||:::
Db        124 IleGlnArgLeuTrpIleValAspSerLeuGluArgLeuGlyIleHisArgHisPheLys 143

Qy        211 AATGAGTTCAAAGAAATCTTGTCCTCTATATATCTCGACCATCACTATTACAAGAACCCT 270
              :::|||    |||      |||      :::|||        ::::::    :::|||
Db        144 AspGluIleLysSerAlaLeuAspTyrValTyr------SerTyrTrpGlyGluAsnGly 161

Qy        271 TTT------------CCAAAAGAAGAAAGGGATCTCTACTCCACATCTCTTGCATTTAGG 318
                                           ||||||    |||||||:::|||       |||
Db        162 IleGlyCysGlyArgGluSerValValThrAspLeuAsnSerThrAlaLeuGlyLeuArg 181

Qy        319 CTCCTCAGAGAACATGGTTTTCAAGTCGCACAAGAGGTATTCGATAGTTTCAAGAACGAG 378
                 ||||||   ||||||:::   |||:::   :::||||||    :::||||||   :::
Db        182 ThrLeuArgLeuHisGlyTyrProValSerSerAspValPheLysAlaPheLysGlyGln 201

Qy        379 GAGGGTGAGTTCAAAGAAAGC------------CTTAGCGACGACACCAGAGGATTGTTG 426
                 |||:::|||      |||                  ::::::    ||||||:::|||
Db        202 AsnGlyGlnPheSerCysSerGluAsnIleGlnThrAspGluGluIleArgGlyValLeu 221

Qy        427 CAACTGTATGAAGCTTCCTTTCTGTTGACGGAAGGCGAAACCACGCTCGAGTCAGCGAGG 486
                 |||:::     ||||||   :::         ||||||       ::::::    |||
Db        222 AsnLeuPheArgAlaSerLeuIleAlaPheProGlyGluLysIleMetAspGluAlaGlu 241

Qy        487 GAATTCGCCACCAAATTTTTGGAGGAAAAAGTGAACGAGGGTGGTGTTGATGGCGACCTT 546
                 |||:::||||||:::|||:::|||    :::    :::       |||       |||
Db        242 IlePheSerThrLysTyrLeuLysGluAlaLeuGlnLysIleProVal---SerSerLeu 260

Qy        547 TTAACAAGAATCGCATATTCTTTGGACATCCCTCTTCATTGGAGGATTAAAAGGCCAAAT 606
                       |||         |||:::          |||       :::    |||
Db        261 SerArgGluIleGlyAspValLeuGluTyrGlyTrpHisThrTyrLeuProArgLeuGlu 280

Qy        607 GCACCTGTGTGGATCGAATGGTATAGGAAGAGGCCCGAC-----------------ATG 648
              |||       :::|||:::   :::    :::
Db        281 AlaArgAsnTyrIleGlnValPheGlyGlnAspThrGluAsnThrLysSerTyrValLys 300

Qy        649 AATCCAGTAGTGTTGGAGCTTGCCATACTCGACTTAAATATTGTTCAAGCACAATTTCAA 708
              :::       :::||||||||||||   |||:::    ||||||   |||:::       :::
Db        301 SerLysLysLeuLeuGluLeuAlaLysLeuGluPheAsnIlePheGlnSerLeuGlnLys 320

Qy        709 GAAGAGCTCAAAGAATCCTTCAGGTGGTGGAGAAATACTGGGTTTGTTGAGAAGCTGCCC 768
                 ||||||:::          |||||||||:::    :::||||||    |||    :::
Db        321 ArgGluLeuGluSerLeuValArgTrpTrpLysGluSerGlyPheProGlu---MetThr 339

Qy        769 TTCGCAAGGGATAGACTGGTGGAATGCTACTTTTGGAATACTGGGATCATCGAGCCACGT 828
              |||   |||    |||    ||||||   |||          :::    |||
Db        340 PheCysArgHisArgHisValGluTyrTyrThrLeuAlaSerCysIleAlaPheGluPro 359

Qy        829 CAGCATGCAAGTGCAAGGATAATGATGGGCAAAGTCAACGCTCTGATTACGGTGATCGAT 888
              ||||||:::          |||:::          |||       ||||||||||||:::|||
Db        360 GlnHisSerGlyPheArgLeuGlyPheAlaLysThrCysHisLeuIleThrValLeuAsp 379

Qy        889 GATATTTATGATGTCTATGGCACCTTAGAAGAACTCGAACAATTCACTGACCTCATTCGA 948
              |||:::||||||   :::||||||::::::||||||||    ||||||       ::::::
Db        380 AspMetTyrAspThrPheGlyThrValAspGluLeuGluLeuPheThrAlaThrMetLys 399

Qy        949 AGATGGGATATAAACTCAATCGACCAACTTCCCGATTACATGCAACTGTGCTTTCTTGCA 1008
              |||||||||      :::|||||||||    ||||||:::||||||:::       ::::::|||
Db        400 ArgTrpAspProSerSerIleAspCysLeuProGluTyrMetLysGlyValTyrIleAla 419

Qy       1009 CTCAACAACTTCGTCGATGATACATCGTACGATGTTATGAAGGAGAAAGGCGTCAACGTT 1068
              :::   :::    |||::::::   :::    :::       :::    :::|||    :::
Db        420 ValTyrAspThrValAsnGluMetAlaArgGluAlaGluGluAlaGlnGlyArgAspThr 439

Qy       1069 ATACCCTACCTGCGGCAATCGTGGGTTGATTTGGCGGATAAGTATATGGTAGAGGCACGG 1128
              :::    |||    |||:::::::|||          |||    ||||||    |||||||||
Db        440 LeuThrTyrAlaArgGluAlaTrpGluAlaTyrIleAspSerTyrMetGlnGluAlaArg 459

Qy       1129 TGGTTCTACGGCGGGCACAAACCAAGTTTGGAAGAGTATTTGGAGAACTCATGGCAGTCG 1188
                 |||          |||:::    ||||||    :::||||||    ||||||       |||
Db        460 TrpIleAlaThrGlyTyrLeuProSerPheAspGluTyrTyrGluAsnGlyLysValSer 479

Qy       1189 ------------ATAAGTGGG------CCCTGTATGTTAACGCACATATTCTTCCGAGTA 1230
                          ||||||         |||    :::          |||    |||
Db        480 CysGlyHisArgIleSerAlaLeuGlnProIleLeuThrMetAspIleProPhe------ 497

Qy       1231 ACAGATTCGTTCACAAAGGAGACCGTCGACAGTTTGTACAAATACCACGATTTAGTTCGT 1290
                 |||          ||||||    ||||||          |||    :::||||||
Db        498 ProAspHisIleLeuLysGlu---ValAspPheProSerLysLeuAsnAspLeu------ 514

Qy       1291 TGGTCATCCTTCGTTCTGCGGCTTGCTGATGATTTGGGAACCTCGGTGGAAGAGGTGAGC 1350
                 :::       :::|||||||||       |||                :::    :::
Db        515 ---AlaCysAlaIleLeuArgLeuArgGlyAspThrArgCysTyrLysAlaAspArgAla 533

Qy       1351 AGAGGGGATGTGCCGAAATCACTTCAGTGCTACATGAGTGACTAC---AATGCATCGGAG 1407
              ||||||:::          |||:::    |||||||||    |||             ||||||
Db        534 ArgGlyGluGluAlaSerSerIleSerCysTyrMetLysAspAsnProGlyValSerGlu 553

Qy       1408 GCGGAGGCGCGGAAGCACGTGAAATGGCTGATAGCGGAGGTGTGGAAGAAGATGAATGCG 1467
                 :::|||    |||:::       :::|||::::::|||    |||    :::|||
Db        554 GluAspAlaLeuAspHisIleAsnAlaMetIleSerAspValIleLysGlyLeuAsnTrp 573

Qy       1468 GAGAGGGTGTCGAAGGATTCTCCATTCGGCAAAGATTTTATAGGATGTGCAGTTGATTTA 1527
                 |||    :::    |||                         |||    |||:::
Db        574 GluLeuLeuLysProAspIleAsnValProIleSerAlaLysLysHisAlaPheAspIle 593

Qy       1528 GGAAGGATGGCGCAGTTGATGTACCATAATGGAGATGGCACGGCACACAACACCCTATT 1587
                 |||          |||       ||||||:::          :::
Db        594 AlaArgAlaPheHisTyrGlyTyrLysTyrArgAspGlyTyrSerValAlaAsnValGlu 613

Qy       1588 ATACATCAACAAATGACCAGAACCTTATTCGAG 1620
                       :::||||||||||||   |||
Db        614 ThrLysSerLeuValThrArgThrLeuLeuGlu 624
```

```
RESULT 8
CASS_RICCO
ID   CASS_RICCO      STANDARD;      PRT;   601 AA.
AC   P59287;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Casbene synthase, chloroplast precursor (EC 4.2.3.8).
OS   Ricinus communis (Castor bean).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   eurosids I; Malpighiales; Euphorbiaceae; Acalyphoideae; Acalypheae;
OC   Ricinus.
OX   NCBI_TaxID=3988;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94359958; PubMed=8078910;
RA   Mau C.J., West C.A.;
RT   "Cloning of casbene synthase cDNA: evidence for conserved structural
RT   features among terpenoid cyclases in plants.";
RL   Proc. Natl. Acad. Sci. U.S.A. 91:8497-8501(1994).
RN   [2]
RP   CHARACTERIZATION.
RX   MEDLINE=97115639; PubMed=8954576;
RA   Hill A.M., Cane D.E., Mau C.J., West C.A.;
RT   "High level expression of Ricinus communis casbene synthase in
RT   Escherichia coli and characterization of the recombinant enzyme.";
RL   Arch. Biochem. Biophys. 336:283-289(1996).
```

```
CC   -!- FUNCTION: Catalyzes the cyclization of geranylgeranyl diphosphate
CC       to casbene, a diterpene phytoalexin with antibacterial and
CC       antifungal activity.
CC   -!- CATALYTIC ACTIVITY: Geranylgeranyl diphosphate = casbene +
CC       diphosphate.
CC   -!- SUBCELLULAR LOCATION: Chloroplast.
CC   -!- INDUCTION: By oligogalacturonide fragments released by fungal
CC       infection. Detected after 5 h of incubation with the pectic
CC       fragments and reaches a maximum after 10-12 h.
CC   -!- MISCELLANEOUS: The Km of this enzyme is 1.9 micromol.
CC   -!- SIMILARITY: Belongs to the terpene synthase family.
CC   -!- CAUTION: It is uncertain whether Met-1 or Met-8 is the initiator.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; L32134; -; NOT_ANNOTATED_CDS.
DR   InterPro; IPR008930; Terp_cyc_toroid.
DR   InterPro; IPR001906; Terp_synth-like.
DR   InterPro; IPR005630; Terpene_synth_C.
DR   InterPro; IPR008949; Terpenoid_synth.
DR   Pfam; PF01397; Terpene_synth; 1.
DR   Pfam; PF03936; Terpene_synth_C; 1.
KW   Plant defense; Lyase; Chloroplast; Transit peptide.
FT   TRANSIT       1     56       CHLOROPLAST (POTENTIAL).
FT   CHAIN        57    601       CASBENE SYNTHASE.
FT   ACT_SITE    500    500       BY SIMILARITY.
FT   ACT_SITE    575    575       BY SIMILARITY.
FT   ACT_SITE    579    579       BY SIMILARITY.
SQ   SEQUENCE   601 AA;  68965 MW;  F7B362D286747957 CRC64;

Alignment Scores:
Pred. No.:              3.65e-43       Length:        601
Score:                  697.00         Matches:       167
Percent Similarity:     51.21%         Conservative:  109
Best Local Similarity:  30.98%         Mismatches:    245
Query Match:            23.97%         Indels:        18
DB:                     1              Gaps:          7

US-09-938-956-6 (1-1632) x CASS_RICCO (1-601)

Qy         19 TACAACCCTTCTCGTTGGGATGTCAACTTCATCCAATCGCTTCTCAGTGACTATAAGGAG 78
Db         69 PheProProThrValTrpGlyAsnArgPheAlaSerLeuThrPheAsnProSerGluPhe 88

Qy         79 GACAAACACGTGATTAGGGCTTCTGAGCTGGTCACTTTGGTGAAG---ATGGAACTGGAG 135
Db         89 GluSerTyrAspGluArgValIleValLeuLysLysLysValLysAspIleLeuIleSer 108

Qy        136 AAAGAAACGGATCAAATTCGACAACTTGAGTTGATCGATGACTTGCAGAGGATGGGGCTG 195
Db        109 SerThrSerAspSerValGluThrValIleLeuIleAspLeuLeuCysArgLeuGlyVal 128

Qy        196 TCCGATCATTTCCAAAATGAGTTCAAAGAAATCTTGTCCTCTATATATCTCGACCATCAC 255
Db        129 SerTyrHisPheGluAsnAspIleGluGluLeuLeuSerLysIlePhe------------ 144

Qy        256 TATTACAAGAACCCTTTTCCA------------AAAGAAGAAAGGGATCTCTACTCCACA 303
Db        145 ---------AsnSerGlnProAspLeuValAspGluLysGluCysAspLeuTyrThrAla 161

Qy        304 TCTCTTGCATTTAGGCTCCTCAGAGAACATGGTTTTCAAGTCGCACAAGAGGTATTCGAT 363
Db        162 AlaIleValPheArgValPheArgGlnHisGlyPheLysMetSerSerAspValPheSer 181

Qy        364 AGTTTCAAGAACGAGGAGGGTGAGTTCAAAGAAAGCCTTAGCGACGACACCAGAGGATTG 423
Db        182 LysPheLysAspSerAspGlyLysPheLysGluSerLeuArgGlyAspAlaLysGlyMet 201

Qy        424 TTGCAACTGTATGAAGCTTCCTTTCTGTTGACGGAAGGCGAAACCACGCTCGAGTCAGCG 483
Db        202 LeuSerLeuPheGluAlaSerHisLeuSerValHisGlyGluAspIleLeuGluGluAla 221

Qy        484 AGGGAATTCGCCACCAAATTTTTGGAGGAAAAAGTGAACGAGGGTGGTGTTGATGGCGAC 543
Db        222 PheAlaPheThrLysAspTyrLeuGlnSerSerAlaValGlu------LeuPheProAsn 239

Qy        544 CTTTTAACAAGAATCGCATATTCTTTGGACATCCCTCTTCATTGGAGGATTAAAAGGCCA 603
Db        240 LeuLysArgHisIleThrAsnAlaLeuGluGlnProPheHisSerGlyValProArgLeu 259

Qy        604 AATGCACCTGTGTGGATCGAATGGTATAGGAAGAGGCCCGAC---ATGAATCCAGTAGTG 660
Db        260 GluAlaArgLysPheIleAspLeuTyrGluAlaAspIleGluCysArgAsnGluThrLeu 279

Qy        661 TTGGAGCTTGCCATACTCGACTTAAATATTGTTCAAGCACAATTTCAAGAAGAGCTCAAA 720
Db        280 LeuGluPheAlaLysLeuAspTyrAsnArgValGlnLeuLeuHisGlnGlnGluLeuCys 299

Qy        721 GAATCCTTCAGGTGGTGGAGAAATACTGGGTTTGTTGAGAAGCTGCCCTTCGCAAGGGAT 780
Db        300 GlnPheSerLysTrpTrpLysAspLeuAsnLeuAlaSerAspIleProTyrAlaArgAsp 319

Qy        781 AGACTGGTGGAATGCTACTTTTGGAATACTGGGATCATCGAGCCACGTCAGCATGCAAGT 840
Db        320 ArgMetAlaGluIlePhePheTrpAlaValAlaMetTyrPheGluProAspTyrAlaHis 339

Qy        841 GCAAGGATAATGATGGGCAAAGTCAACGCTCTGATTACGGTGATCGATGATATTTATGAT 900
Db        340 ThrArgMetIleIleAlaLysValValLeuLeuIleSerLeuIleAspAspThrIleAsp 359

Qy        901 GTCTATGGCACCTTAGAAGAACTCGAACAATTCACTGACCTCATTCGAAGATGGGATATA 960
Db        360 AlaTyrAlaThrMetGluGluThrHisIleLeuAlaGluAlaValAlaArgTrpAspMet 379

Qy        961 AACTCAATCGACCAACTTCCCGATTACATGCAACTGTGCTTTCTTGCACTCAACAACTTC 1020
Db        380 SerCysLeuGluLysLeuProAspTyrMetLysValIleTyrLysLeuLeuLeuAsnThr 399

Qy       1021 GTCGATGATACATCGTACGATGTTATGAAGGAGAAAGGCGTCAACGTTATACCCTACCTG 1080
Db        400 PheSerGluPheGluLysGluLeuThrAlaGluGlyLysSerTyrSerValLysTyrGly 419

Qy       1081 CGGCAATCGTGGGTTGATTTGGCGGATAAGTATATGGTAGAGGCACGGTGGTTCTACGGC 1140
Db        420 ArgGluAlaPheGlnGluLeuValArgGlyTyrTyrLeuGluAlaValTrpArgAspGlu 439

Qy       1141 GGGCACAAACCAAGTTTGGAAGAGTATTTGGAGAACTCATGGCAGTCGATAAGTGGGCCC 1200
Db        440 GlyLysIleProSerPheAspAspTyrLeuTyrAsnGlySerMetThrThrGlyLeuPro 459

Qy       1201 TGTATGTTAACGCACATATTCTTCCGAGTAACAGATTCGTTCACAAAGGAGACCGTCGAC 1260
Db        460 LeuValSerThrAlaSerPheMetGlyValGlnGluIleThrGlyLeuAsnGluPheGln 479

Qy       1261 AGTTTGTACAAATACCACGATTTAGTTCGTTGGTCATCCTTCGTTCTGCGGCTTGCTGAT 1320
Db        480 TrpLeuGluThrAsnProLysLeuSerTyrAlaSerGlyAlaPheIleArgLeuValAsn 499

Qy       1321 GATTTGGGAACCTCGGTGGAAGAGGTGAGCAGAGGGGATGTGCCGAAATCACTTCAGTGC 1380
Db        500 AspLeuThrSerHisValThrGluGlnGlnArgGlyHisValAlaSerCysIleAspCys 519

Qy       1381 TACATGAGTGACTACAATGCATCGGAGGCGGAGGCGCGGAAGCACGTGAAATGGCTGATA 1440
Db        520 TyrMetAsnGlnHisGlyValSerLysAspGluAlaValLysIleLeuGlnLysMetAla 539

Qy       1441 GCGGAGGTGTGGAAGAAGATGAATGCGGAGAGGGTGTCGAAGGATTCTCCATTCGGCAAA 1500
Db        540 ThrAspCysTrpLysGluIleAsn---GluGluCysMetArgGlnSerGlnValSerVal 558
```

```
Qy       1501 GATTTTATAGGATGTGCAGTTGATTTAGGAAGGATGGCGCAGTTGATGTACCATAATGGA 1560
                   :::         |||:::|||  |||:::       :::    |||       |||
Db        559 GlyHisLeuMetArgIleValAsnLeuAlaArgLeuThrAspValSerTyrLysTyrGly 578

Qy       1561 GATGGGCACGGCACACAACACCCTATTATACATCAACAAATGACCAGAACCTTATTC 1617
                 ||||||:::               :::      :::|||     :::    ||||||
Db        579 AspGlyTyrThrAspSerGlnGlnLeu------LysGlnPheValLysGlyLeuPhe 595

RESULT 9
TSD3_ABIGR
ID   TSD3_ABIGR     STANDARD;      PRT;   637 AA.
AC   O22340;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   (-)-(4S)-limonene synthase, chloroplast precursor (EC 4.2.3.16).
GN   AG10.
OS   Abies grandis (Grand fir).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Abies.
OX   NCBI_TaxID=46611;
RN   [1]
RP   SEQUENCE FROM N.A., AND CHARACTERIZATION.
RX   MEDLINE=97413772; PubMed=9268308;
RA   Bohlmann J., Steele C.L., Croteau R.;
RT   "Monoterpene synthases from grand fir (Abies grandis). cDNA isolation,
RT   characterization, and functional expression of myrcene synthase, (-)-
RT   (4S)-limonene synthase, and (-)-(1S,5S)-pinene synthase.";
RL   J. Biol. Chem. 272:21784-21792(1997).
CC   -!- FUNCTION: Involved in defensive oleoresin formation in conifers in
CC       response to insect attack or other injury. Involved in monoterpene
CC       (C10) olefins biosynthesis.
CC   -!- CATALYTIC ACTIVITY: Geranyl diphosphate = limonene + diphosphate.
CC   -!- COFACTOR: Manganese and potassium.
CC   -!- PATHWAY: Oleoresinosis.
CC   -!- SUBCELLULAR LOCATION: Chloroplast.
CC   -!- INDUCTION: By wounding.
CC   -!- SIMILARITY: Belongs to the terpene synthase family.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AF006193; AAB70907.1; -.
DR   InterPro; IPR008930; Terp_cyc_toroid.
DR   InterPro; IPR001906; Terp_synth-like.
DR   InterPro; IPR005630; Terpene_synth_C.
DR   InterPro; IPR008949; Terpenoid_synth.
DR   Pfam; PF01397; Terpene_synth; 1.
DR   Pfam; PF03936; Terpene_synth_C; 1.
KW   Lyase; Manganese; Transit peptide; Chloroplast.
FT   TRANSIT       1      ?       CHLOROPLAST (POTENTIAL).
FT   CHAIN         ?    637       (-)-(4S)-LIMONENE SYNTHASE.
FT   ACT_SITE    533    533       BY SIMILARITY.
FT   ACT_SITE    610    610       BY SIMILARITY.
FT   ACT_SITE    614    614       BY SIMILARITY.
SQ   SEQUENCE   637 AA;  73477 MW;  8E80CD9DDE886898 CRC64;

Alignment Scores:
Pred. No.:                4.99e-42     Length:        637
Score:                    681.50       Matches:       181
Percent Similarity:       51.46%       Conservative:  119
Best Local Similarity:    31.05%       Mismatches:    226
Query Match:              23.44%       Indels:        57
DB:                       1            Gaps:          16

US-09-938-956-6 (1-1632) x TSD3_ABIGR (1-637)

Qy          4 AGACGATCCGGAAACTACAACCCTTCTCGTTGGGATGTCAACTTCATCCAATCGCTTCTC 63
              ||||||         :::::::::|||:::    |||:::   :::||||||||||||||
Db         70 ArgArgIleAlaAspHisHisProAsnLeuTrpGluAspAspPheIleGlnSerLeuSer 89

Qy         64 AGTGACTATAAGGAGGACAAACACGTGATTAGGGCTTCTGAGCTGGTCACTTTGGTGAAG 123
              |||   |||                :::        ||||||      :::|||    ||||||
Db         90 SerProTyrGlyGlySerSerTyrSerGluArgAlaGluThrValValGluGluValLys 109

Qy        124 --------------------------ATGGAACTG---GAGAAAGAAACGGATCAAATT 153
                                           ||||||      :::   |||   :::
Db        110 GluMetPheAsnSerIleProAsnAsnArgGluLeuPheGlySerGlnAsnAspLeuLeu 129

Qy        154 CGACAACTTGAGTTGATCGATGACTTGCAGAGGATGGGGCTGTCCGATCATTTCCAAAAT 213
                 :::|||   ::::::|||   ::::::|||:::|||:::        |||||||||||
Db        130 ThrArgLeuTrpMetValAspSerIleGluArgLeuGlyIleAspArgHisPheGlnAsn 149

Qy        214 GAGTTCAAAGAAATCTTGTCCTCTATATATCTCGACCATCACTATTACAAG--------- 264
              |||   :::         |||      :::|||         |||:::|||
Db        150 GluIleArgValAlaLeuAspTyrValTyr---------SerTyrTrpLysGluLysGlu 166

Qy        265 ---------------------AACCCTTTTCCAAAAGAAGAAAGGGATCTCTACTCCACA 303
                                   :::   ||||||            ||||||   ||||||
Db        167 GlyIleGlyCysGlyArgAspSerThrPhePro------------AspLeuAsnSerThr 182

Qy        304 TCTCTTGCATTTAGGCTCCTCAGAGAACATGGTTTTCAAGTCGCACAAGAGGTATTCGAT 363
              :::||||||   |||   ||||||   ||||||:::   |||:::   :::|||   :::
Db        183 AlaLeuAlaLeuArgThrLeuArgLeuHisGlyTyrAsnValSerSerAspValLeuGlu 202

Qy        364 AGTTTCAAGAACGAGGAGGGTGAGTTC------AAAGAAAGCCTTAGCGACGAC------ 411
              ||||||:::|||:::|||          |||              |||::::::
Db        203 TyrPheLysAspGluLysGlyHisPheAlaCysProAlaIleLeuThrGluGlyGlnIle 222

Qy        412 ACCAGAGGATTGTTGCAACTGTATGAAGCTTCCTTTCTGTTGACGGAAGGCGAAACCACG 471
              ||||||   :::|||   ||||||   ||||||   :::         ||||||
Db        223 ThrArgSerValLeuAsnLeuTyrArgAlaSerLeuValAlaPheProGlyGluLysVal 242

Qy        472 CTCGAGTCAGCGAGGGAATTCGCCACCAAATTTTTGGAG---------GAAAAAGTGAAC 522
              :::|||   |||      |||:::      :::|||:::         :::|||:::
Db        243 MetGluGluAlaGluIlePheSerAlaSerTyrLeuLysLysValLeuGlnLysIlePro 262

Qy        523 GAGGGTGGTGTTGATGGCGACCTTTTAACAAGAATCGCATATTCTTTGGACATCCCTCTT 582
                    :::   |||:::            |||   |||   |||:::
Db        263 ValSerAsnLeuSerGlyGlu------------IleGluTyrValLeuGluTyrGlyTrp 278

Qy        583 CATTGGAGGATTAAAAGGCCAAATGCACCTGTGTGGATCGAATGGTAT------------ 630
              |||   :::   |||      |||      :::||||||   |||
Db        279 HisThrAsnLeuProArgLeuGluAlaArgAsnTyrIleGluValTyrGluGlnSerGly 298

Qy        631 ------------AGGAAGAGGCCCGACATGAAT---CCAGTAGTGTTGGAGCTTGCCATA 675
                          :::   |||   ||||||         :::|||:::||||||
Db        299 TyrGluSerLeuAsnGluMetProTyrMetAsnMetLysLysLeuLeuGlnLeuAlaLys 318

Qy        676 CTCGACTTAAATATTGTTCAAGCACAATTTCAAGAAGAGCTCAAAGAATCCTTCAGGTGG 735
              |||:::   ||||||      :::         ||||||:::         ||||||
Db        319 LeuGluPheAsnIlePheHisSerLeuGlnLeuArgGluLeuGlnSerIleSerArgTrp 338

Qy        736 TGGAGAAATACTGGGTTTGTTGAGAAGCTGCCCTTCGCAAGGGATAGACTGGTGGAATGC 795
              |||:::   :::|||         :::|||   |||   |||   |||   ||||||
Db        339 TrpLysGluSerGly---SerSerGlnLeuThrPheThrArgHisArgHisValGluTyr 357

Qy        796 TACTTTTGGAATACTGGGATCATCGAGCCACGTCAGCATGCAAGTGCAAGGATAATGATG 855
              |||      :::   |||         :::|||::::::      |||:::
Db        358 TyrThrMetAlaSerCysIleSerMetLeuProLysHisSerAlaPheArgMetGluPhe 377

Qy        856 GGCAAAGTCAACGCTCTGATTACGGTGATCGATGATATTTATGATGTCTATGGCACCTTA 915
                 ||||||      |||:::|||||||||:::||||||||||||||      :::|||||||:::
Db        378 ValLysValCysHisLeuValThrValLeuAspAspIleTyrAspThrPheGlyThrMet 397

Qy        916 GAAGAACTCGAACAATTCACTGACCTCATTCGAAGATGGGATATAAACTCAATCGACCAA 975
                 ||||||:::   |||||||||   |||:::|||||||||:::::::::
Db        398 AsnGluLeuGlnLeuPheThrAspAlaIleLysArgTrpAspLeuSerThrThrArgTrp 417

Qy        976 CTTCCCGATTACATGCAACTGTGCTTTCTTGCACTCAACAACTTCGTCGATGATACATCG 1035
              ||||||:::||||||:::      ::::::   |||         :::::::::
```

```
Db        418 LeuProGluTyrMetLysGlyValTyrMetAspLeuTyrGlnCysIleAsnGluMetVal 437
Qy       1036 TACGATGTTATGAAGGAGAAAGGCGTCAACGTTATACCCTACCTGCGGCAATCGTGGGTT 1095
Db        438 GluGluAlaGluLysThrGlnGlyArgAspMetLeuAsnTyrIleGlnAsnAlaTrpGlu 457
Qy       1096 GATTTGGCGGATAAGTATATGGTAGAGGCACGGTGGTTCTACGGCGGGCACAAACCAAGT 1155
Db        458 AlaLeuPheAspThrPheMetGlnGluAlaLysTrpIleSerSerSerTyrLeuProThr 477
Qy       1156 TTGGAAGAGTATTTGGAGAACTCATGGCAGTCGATAAGTGGGCCCTGTATGTTAACGCAC 1215
Db        478 PheGluGluTyrLeuLysAsnAlaLysValSer---SerGlySerArgIleAlaThrLeu 496
Qy       1216 ATATTCTTCCGAGTAACAGAT------------TCGTTCACAAAGGAGACCGTCGACAGT 1263
Db        497 GlnProIleLeuThrLeuAspValProLeuProAspTyrIleLeuGlnGluIleAspTyr 516
Qy       1264 TTGTACAAATACCACGATTTAGTTCGTTGGTCATCCTTCGTTCTGCGGCTTGCTGATGAT 1323
Db        517 ProSerArgPheAsnGluLeu---------AlaSerSerIleLeuArgLeuArgGlyAsp 533
Qy       1324 TTGGGAACCTCGGTGGAAGAGGTGAGCAGAGGGGATGTGCCGAAATCACTTCAGTGCTAC 1383
Db        534 ThrArgCysTyrLysAlaAspArgAlaArgGlyGluGluAlaSerAlaIleSerCysTyr 553
Qy       1384 ATGAGTGACTACAATGCATCG---GAGGCGGAGGCGCGGAAGCACGTGAAATGGCTGATA 1440
Db        554 MetLysAspHisProGlySerIleGluGluAspAlaLeuAsnHisIleAsnAlaMetIle 573
Qy       1441 GCGGAGGTGTGGAAGAAGATGAATGCGGAGAGGGTGTCGAAGGATTCTCCATTCGGCAAA 1500
Db        574 SerAspAlaIleArgGluLeuAsnTrpGluLeuLeuArgProAspSerLysSerProIle 593
Qy       1501 GATTTTATAGGATGTGCAGTTGATTTAGGAAGGATGGCGCAGTTGATGTACCATAATGGA 1560
Db        594 SerSerLysLysHisAlaPheAspIleThrArgAlaPheHisHisValTyrLysTyrArg 613
Qy       1561 GATGGGCACGGCACACAACACCCTATTATACATCAACAAATGACCAGAACCTTATTCGAG 1620
Db        614 AspGlyTyrThrValSerAsnAsnGluThrLysAsnLeuValMetLysThrValLeuGlu 633
Qy       1621 CCCTTTGCA 1629
Db        634 ProLeuAla 636

RESULT 10
TSD2_ABIGR
ID   TSD2_ABIGR     STANDARD;      PRT;   627 AA.
AC   O24474;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Myrcene synthase, chloroplast precursor (EC 4.2.3.15).
GN   AG2.2.
OS   Abies grandis (Grand fir).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Abies.
OX   NCBI_TaxID=46611;
RN   [1]
RP   SEQUENCE FROM N.A., AND CHARACTERIZATION.
RX   MEDLINE=97413772; PubMed=9268308;
RA   Bohlmann J., Steele C.L., Croteau R.;
RT   "Monoterpene synthases from grand fir (Abies grandis). cDNA isolation,
RT   characterization, and functional expression of myrcene synthase, (-)-
RT   (4S)-limonene synthase, and (-)-(1S,5S)-pinene synthase.";
RL   J. Biol. Chem. 272:21784-21792(1997).
CC   -!- FUNCTION: Involved in defensive oleoresin formation in conifers in
CC       response to insect attack or other injury. Involved in monoterpene
CC       (C10) olefins biosynthesis.
CC   -!- CATALYTIC ACTIVITY: Geranyl diphosphate = myrcene + diphosphate.
CC   -!- COFACTOR: Manganese and potassium.
CC   -!- PATHWAY: Oleoresinosis.
CC   -!- SUBCELLULAR LOCATION: Chloroplast.
CC   -!- INDUCTION: By wounding.
CC   -!- SIMILARITY: Belongs to the terpene synthase family.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; U87908; AAB71084.1; -.
DR   HSSP; Q40577; 5EAS.
DR   InterPro; IPR008930; Terp_cyc_toroid.
DR   InterPro; IPR001906; Terp_synth-like.
DR   InterPro; IPR005630; Terpene_synth_C.
DR   InterPro; IPR008949; Terpenoid_synth.
DR   Pfam; PF01397; Terpene_synth; 1.
DR   Pfam; PF03936; Terpene_synth_C; 1.
KW   Lyase; Manganese; Transit peptide; Chloroplast.
FT   TRANSIT       1      ?       CHLOROPLAST (POTENTIAL).
FT   CHAIN         ?    627       MYRCENE SYNTHASE.
FT   ACT_SITE    523    523       BY SIMILARITY.
FT   ACT_SITE    600    600       BY SIMILARITY.
FT   ACT_SITE    604    604       BY SIMILARITY.
SQ   SEQUENCE   627 AA;  72478 MW;  2E0DA492E0C971FD CRC64;

Alignment Scores:
Pred. No.:               2.26e-41      Length:         627
Score:                   672.50        Matches:        177
Percent Similarity:      49.74%        Conservative:   114
Best Local Similarity:   30.26%        Mismatches:     223
Query Match:             23.13%        Indels:         71
DB:                      1             Gaps:           15

US-09-938-956-6 (1-1632) x TSD2_ABIGR (1-627)

Qy          4 AGACGATCCGGAAACTACAACCCTTCTCGTTGGGATGTCAACTTCATCCAATCGCTTCTC 63
Db         64 ArgArgIleGlyAspTyrHisSerAsnIleTrpAspAspAspPheIleGlnSerLeuSer 83
Qy         64 AGTGACTATAAGGAGGACAAACACGTGATTAGGGCTTCTGAGCTGGTCACTTTGGTGAAG 123
Db         84 ThrProTyrGlyGluProSerTyrGlnGluArgAlaGluArgLeuIleValGluValLys 103
Qy        124 ---------------ATGGAACTGGAG-----------------AAAGAAACGGATCAA 150
Db        104 LysIlePheAsnSerMetTyrLeuAspAspGlyArgLeuMetSerSerPheAsnAspLeu 123
Qy        151 ATTCGACAACTTGAGTTGATCGATGACTTGCAGAGGATGGGGCTGTCCGATCATTTCCAA 210
Db        124 MetGlnArgLeuTrpIleValAspSerValGluArgLeuGlyIleAlaArgHisPheLys 143
Qy        211 AATGAGTTCAAAGAAATCTTGTCCTCTATATATCTCGACCATCACTATTACAAGAACCCT 270
Db        144 AsnGluIleThrSerAlaLeuAspTyrValPhe---------ArgTyrTrp--------- 157
Qy        271 TTTCCAAAAGAAGAAAGG--------------------------------GATCTCTAC 297
Db        158 ---------GluGluAsnGlyIleGlyCysGlyArgAspSerIleValThrAspLeuAsn 174
Qy        298 TCCACATCTCTTGCATTTAGGCTCCTCAGAGAACATGGTTTTCAAGTCGCACAAGAGGTA 357
Db        175 SerThrAlaLeuGlyPheArgThrLeuArgLeuHisGlyTyrThrValSerProGluVal 194
Qy        358 TTCGATAGTTTCAAGAACGAGGAGGGTGAGTTCAAAGAAAGC---------CTTAGCGAC 408
Db        195 LeuLysAlaPheGlnAspGlnAsnGlyGlnPheValCysSerProGlyGlnThrGluGly 214
Qy        409 GACACCAGAGGATTGTTGCAACTGTATGAAGCTTCCTTTCTGTTGACGGAAGGCGAAACC 468
Db        215 GluIleArgSerValLeuAsnLeuTyrArgAlaSerLeuIleAlaPheProGlyGluLys 234
```

```
Qy    469 ACGCTCGAGTCAGCGAGGGAATTCGCCACCAAATTTTTGGAGGAAAAAGTGAACGAGGGT 528
           :::|||  |||   |||:::|||::::::|||:::|||   :::   :::
Db    235 ValMetGluGluAlaGluIlePheSerThrArgTyrLeuLysGluAlaLeuGlnLysIle 254

Qy    529 GGTGTTGATGGCGACCTTTTAACAAGAATCGCATATTCTTTGGACATCCCTCTTCATTGG 588
             |||      |||      |||   :::  ::::::           |||
Db    255 ProValSerAla---LeuSerGlnGluIleLysPheValMetGluTyrGlyTrpHisThr 273

Qy    589 AGGATTAAAAGGCCAAATGCACCTGTGTGGATCGAA------------------------ 624
           :::   |||   |||    :::|||:::
Db    274 AsnLeuProArgLeuGluAlaArgAsnTyrIleAspThrLeuGluLysAspThrSerAla 293

Qy    625 TGGTATAGGAAGAGGCCCGACATGAATCCAGTAGTGTTGGAGCTTGCCATACTCGACTTA 684
          |||   |||      :::         :::||||||||||||   |||:::
Db    294 TrpLeuAsnLys------AsnAlaGlyLysLysLeuLeuGluLeuAlaLysLeuGluPhe 311

Qy    685 AATATTGTTCAAGCACAATTTCAAGAAGAGCTCAAAGAATCCTTCAGGTGGTGGAGAAAT 744
          ||||||     :::       |||:::||||||:::            |||||||||:::
Db    312 AsnIlePheAsnSerLeuGlnGlnLysGluLeuGlnTyrLeuLeuArgTrpTrpLysGlu 331

Qy    745 ACTGGGTTTGTTGAGAAGCTGCCCTTCGCAAGGGATAGACTGGTGGAATGCTACTTTTGG 804
          :::      :::   ||||||  |||||||||   |||   ||||||   |||
Db    332 SerAsp---LeuProLysLeuThrPheAlaArgHisArgHisValGluPheTyrThrLeu 350

Qy    805 AATACTGGGATCATCGAGCCACGTCAGCATGCAAGTGCAAGGATAATGATGGGCAAAGTC 864
           :::   |||         :::|||::::::   |||:::         |||:::
Db    351 AlaSerCysIleAlaIleAspProLysHisSerAlaPheArgLeuGlyPheAlaLysMet 370

Qy    865 AACGCTCTGATTACGGTGATCGATGATATTTATGATGTCTATGGCACCTTAGAAGAACTC 924
             |||:::||||||:::||||||||||||||     :::||||||::::::||||||
Db    371 CysHisLeuValThrValLeuAspAspIleTyrAspThrPheGlyThrIleAspGluLeu 390

Qy    925 GAACAATTCACTGACCTCATTCGAAGATGGGATATAAACTCAATCGACCAACTTCCCGAT 984
          |||  ||||||   |||:::||||||   :::   |||:::   ||||||:::
Db    391 GluLeuPheThrSerAlaIleLysArgTrpAsnSerSerGluIleGluHisLeuProGlu 410

Qy    985 TACATGCAACTGTGCTTTCTTGCACTCAACAACTTCGTCGATGATACATCGTACGATGTT 1044
          ||||||:::       ::::::   :::      |||::::::   :::   :::
Db    411 TyrMetLysCysValTyrMetValValPheGluThrValAsnGluLeuThrArgGluAla 430

Qy   1045 ATGAAGGAGAAAGGCGTCAACGTTATACCCTACCTGCGGCAATCGTGGGTTGATTTGGCG 1104
             |||   :::|||   |||   :::   |||:::|||::::::|||
Db    431 GluLysThrGlnGlyArgAsnThrLeuAsnTyrValArgLysAlaTrpGluAlaTyrPhe 450

Qy   1105 GATAAGTATATGGTAGAGGCACGGTGGTTCTACGGCGGGCACAAACCAAGTTTGGAAGAG 1164
          |||   ||||||   ||||||:::|||         |||:::   |||      ||||||
Db    451 AspSerTyrMetGluGluAlaLysTrpIleSerAsnGlyTyrLeuProMetPheGluGlu 470

Qy   1165 TATTTGGAGAACTCATGGCAGTCGATAAGTGGGCCCTGTATGTTAACGCACATATTCTTC 1224
          |||   ||||||                                             :::
Db    471 TyrHisGluAsnGlyLysValSerSerAla---------------------------Tyr 481

Qy   1225 CGAGTAACA-----------------------------------GATTCGTTCACAAAG 1248
          ||||||                                             :::
Db    482 ArgValAlaThrLeuGlnProIleLeuThrLeuAsnAlaTrpLeuProAspTyrIleLeu 501

Qy   1249 GAGACCGTCGACAGTTTGTACAAATACCACGATTTAGTTCGTTGGTCATCCTTCGTTCTG 1308
          :::   :::|||         :::::::::||||||         |||||||||   |||
Db    502 LysGlyIleAspPheProSerArgPheAsnAspLeuAla------SerSerPhe---Leu 518

Qy   1309 CGGCTTGCTGATGATTTGGGAACCTCGGTGGAAGAGGTGAGCAGAGGGGATGTGCCGAAA 1368
          ||||||      |||                 :::      ||||||:::
Db    519 ArgLeuArgGlyAspThrArgCysTyrLysAlaAspArgAspArgGlyGluGluAlaSer 538

Qy   1369 TCACTTCAGTGCTACATGAGTGACTAC---AATGCATCGGAGGCGGAGGCGCGGAAGCAC 1425
             :::   |||||||||   |||      ::::::|||   :::|||      |||
Db    539 CysIleSerCysTyrMetLysAspAsnProGlySerThrGluGluAspAlaLeuAsnHis 558

Qy   1426 GTGAAATGGCTGATAGCGGAGGTGTGGAAGAAGATGAATGCGGAGAGGGTGTCGAAGGAT 1485
           :::      ::::::   ::::::   |||::::::|||   |||   :::      :::
Db    559 IleAsnAlaMetValAsnAspIleIleLysGluLeuAsnTrpGluLeuLeuArgSerAsn 578

Qy   1486 TCTCCATTCGGCAAAGATTTTATAGGATGTGCAGTTGATTTAGGAAGGATGGCGCAGTTG 1545
                                           |||   |||:::   |||
Db    579 AspAsnIleProMetLeuAlaLysLysHisAlaPheAspIleThrArgAlaLeuHisHis 598

Qy   1546 ATGTACCATAATGGAGATGGGCACGGCACACAACACCCTATTATACATCAACAAATGACC 1605
           :::|||         ||||||          :::            :::   :::
Db    599 LeuTyrIleTyrArgAspGlyPheSerValAlaAsnLysGluThrLysLysLeuValMet 618

Qy   1606 AGAACCTTATTCGAG 1620
              ||||||   |||
Db    619 GluThrLeuLeuGlu 623
```

```
RESULT 11
TASY_TAXBR
ID   TASY_TAXBR     STANDARD;      PRT;   862 AA.
AC   Q41594; Q94FV8;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Taxadiene synthase (EC 4.2.3.17) (Taxa-4(5),11(12)-diene synthase).
GN   TDC1.
OS   Taxus brevifolia (Pacific yew).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Coniferopsida; Coniferales; Taxaceae; Taxus.
OX   NCBI_TaxID=46220;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96199163; PubMed=8621577;
RA   Wildung M.R., Croteau R.B.;
RT   "A cDNA clone for taxadiene synthase, the diterpene cyclase that
RT   catalyzes the committed step of taxol biosynthesis.";
RL   J. Biol. Chem. 271:9201-9204(1996).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21297238; PubMed=11404343;
RA   Trapp S.C., Croteau R.B.;
RT   "Genomic organization of plant terpene synthases and molecular
RT   evolutionary implications.";
RL   Genetics 158:811-832(2001).
CC   -!- FUNCTION: Catalyzes the cyclization of the ubiquitous isoprenoid
CC       intermediate geranylgeranyl diphosphate to taxa-4,11-diene, the
CC       parent olefin with a taxane skeleton.
CC   -!- CATALYTIC ACTIVITY: Geranylgeranyl diphosphate = taxa-4,11-diene +
CC       diphosphate.
CC   -!- PATHWAY: Taxol biosynthesis; first step.
CC   -!- SIMILARITY: Belongs to the terpene synthase family.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; U48796; AAC49310.1; -.
DR   EMBL; AF326519; AAK83566.1; -.
DR   HSSP; Q40577; 5EAU.
DR   InterPro; IPR008930; Terp_cyc_toroid.
DR   InterPro; IPR001906; Terp_synth-like.
DR   InterPro; IPR005630; Terpene_synth_C.
DR   InterPro; IPR008949; Terpenoid_synth.
DR   Pfam; PF01397; Terpene_synth; 1.
DR   Pfam; PF03936; Terpene_synth_C; 1.
KW   Taxol biosynthesis; Lyase.
FT   ACT_SITE    758    758       BY SIMILARITY.
FT   ACT_SITE    835    835       BY SIMILARITY.
FT   ACT_SITE    839    839       BY SIMILARITY.
FT   CONFLICT    148    148       L -> V (IN REF. 2).
FT   CONFLICT    767    767       A -> V (IN REF. 2).
SQ   SEQUENCE   862 AA;  98303 MW;  9141B59780CD79A1 CRC64;

Alignment Scores:
```

```
Pred. No.:                 7.95e-39       Length:          862
Score:                     638.00         Matches:         161
Percent Similarity:        50.86%         Conservative:    106
Best Local Similarity:     30.67%         Mismatches:      210
Query Match:               21.94%         Indels:          48
DB:                        1              Gaps:            17

US-09-938-956-6 (1-1632) x TASY_TAXBR (1-862)

Qy     145 GATCAAATTCGACAACTTGAGTTGATCGATGACTTGCAGAGGATGGGGCTGTCCGATCAT 204
           |||   :::   :::|||  |||:::|||:::::::::   :::|||:::        |||
Db     351 AspLeuLeuGluArgLeuSerLeuValAspAsnIleGluHisLeuGlyIleGlyArgHis 370

Qy     205 TTCCAAAATGAGTTCAAAGAAATCTTGTCCTCTATATATCTCGACCATCACTATTACAAG 264
           |||:::   |||   |||      |||      :::|||           |||:::
Db     371 PheLysGlnGluIleLysGlyAlaLeuAspTyrValTyr------ArgHisTrp------ 386

Qy     265 AACCCTTTTCCAAAAGAAGAAAGG--------------------------------GAT 291
                                   |||||                            |||
Db     387 ---------------SerGluArgGlyIleGlyTrpGlyArgAspSerLeuValProAsp 401

Qy     292 CTCTACTCCACATCTCTTGCATTTAGGCTCCTCAGAGAACATGGTTTTCAAGTCGCACAA 351
           |||   :::|||:::|||      |||   |||||   ||||||:::   |||:::
Db     402 LeuAsnThrThrAlaLeuGlyLeuArgThrLeuArgMetHisGlyTyrAsnValSerSer 421

Qy     352 GAGGTATTCGATAGTTTCAAGAACGAGGAGGGTGAGTTCAAAGAAAGCCTTAGCGAC--- 408
           :::|||   ::::::|||||:::|||   |||      |||
Db     422 AspValLeuAsnAsnPheLysAspGluAsnGlyArgPhePheSerSerAlaGlyGlnThr 441

Qy     409 ------GACACCAGAGGATTGTTGCAACTGTATGAAGCTTCCTTTCTGTTGACGGAAGGC 462
                 :::   |||   ::::::   |||:::   ||||||   |||
Db     442 HisValGluLeuArgSerValValAsnLeuPheArgAlaSerAspLeuAlaPheProAsp 461

Qy     463 GAAACCACGCTCGAGTCAGCGAGGGAATTCGCCACCAAATTTTTGGAGGAAAAAGTG--- 519
           |||      ::::::   ||||||:::||||||      :::|||   |||   :::
Db     462 GluArgAlaMetAspAspAlaArgLysPheAlaGluProTyrLeuArgGluAlaLeuAla 481

Qy     520 AACGAGGGTGGTGTTGATGGCGACCTTTTAACAAGAATCGCATATTCTTTGGACATCCCT 579
              :::      :::      |||         |||   |||   ::::::   |||
Db     482 ThrLysIleSerThrAsnThrLysLeuPheLysGluIleGluTyrValValGluTyrPro 501

Qy     580 CTTCATTGGAGGATTAAAAGGCCAAATGCACCTGTGTGGATCGAA--------------- 624
              |||   |||   |||      |||   :::|||:::
Db     502 TrpHisMetSerIleProArgLeuGluAlaArgSerTyrIleAspSerTyrAspAspAsn 521

Qy     625 ------TGGTATAGGAAG------------AGGCCCGACATG---AATCCAGTAGTGTTG 663
                 |||   ||||||            |||   :::   |||         |||
Db     522 TyrValTrpGlnArgLysThrLeuTyrArgMetProSerLeuSerAsnSerLysCysLeu 541

Qy     664 GAGCTTGCCATACTCGACTTAAATATTGTTCAAGCACAATTTCAAGAAGAGCTCAAAGAA 723
           ||||||||   ||||||   ||||||||||||:::      ||||||||||||||
Db     542 GluLeuAlaLysLeuAspPheAsnIleValGlnSerLeuHisGlnGluGluLeuLysLeu 561

Qy     724 TCCTTCAGGTGGTGGAGAAATACTGGGTTTGTTGAGAAGCTGCCCTTCGCAAGGGATAGA 783
              |||||||||:::   :::|||   :::   |||   |||   |||   |||
Db     562 LeuThrArgTrpTrpLysGluSerGlyMetAlaAsp---IleAsnPheThrArgHisArg 580

Qy     784 CTGGTGGAATGCTACTTTTGGAATACTGGGATCATCGAGCCACGTCAGCATGCAAGTGCA 843
           :::   |||   ||||||   ::::::         ||||||   ::::::::::::
Db     581 ValAlaGluValTyrPhe---SerSerAlaThrPheGluPro---GluTyrSerAlaThr 598

Qy     844 AGGATAATGATGGGCAAAGTCAACGCTCTGATTACGGTGATCGATGATATTTATGATGTC 903
           ||||||         |||:::      |||      :::   ||||||:::   |||:::
Db     599 ArgIleAlaPheThrLysIleGlyCysLeuGlnValLeuPheAspAspMetAlaAspIle 618

Qy     904 TATGGCACCTTAGAAGAACTCGAACAATTCACTGACCTCATTCGAAGATGGGATATAAAC 963
           :::   ||||||:::||||||:::   ||||||:::   ::::::|||||||||   :::
Db     619 PheAlaThrLeuAspGluLeuLysSerPheThrGluGlyValLysArgTrpAspThrSer 638

Qy     964 TCAATCGACCAACTTCCCGATTACATGCAACTGTGCTTTCTTGCACTCAACAACTTCGTC 1023
              :::   ::::::|||:::   ||||||   ||||||               :::
Db     639 LeuLeuHisGluIleProGluCysMetGlnThrCysPheLysValTrpPheLysLeuMet 658

Qy    1024 GATGATACATCGTACGATGTTATGAAGGAGAAAGGCGTCAACGTTATACCCTACCTGCGG 1083
           ::::::   :::   ||||||:::|||   :::|||   :::::::::   ::::::|||
Db     659 GluGluValAsnAsnAspValValLysValGlnGlyArgAspMetLeuAlaHisIleArg 678

Qy    1084 CAATCGTGGGTTGATTTGGCGGATAAGTATATGGTAGAGGCACGGTGGTTCTACGGCGGG 1143
           :::   |||         :::   |||:::   |||      |||         |||
Db     679 LysProTrpGluLeuTyrPheAsnCysTyrValGlnGluArgGluTrpLeuGluAlaGly 698

Qy    1144 CACAAACCAAGTTTGGAAGAGTATTTGGAGAACTCATGGCAGTCGATAAGT---GGGCCC 1200
           :::   |||:::   ||||||||||||:::            |||:::      ||||||
Db     699 TyrIleProThrPheGluGluTyrLeuLysThrTyrAlaIleSerValGlyLeuGlyPro 718

Qy    1201 TGTATGTTAACGCACATATTCTTCCGAGTAACAGATTCGTTCACAAAGGAGACCGTCGAC 1260
           |||   |||      |||         :::   :::            :::   |||:::
Db     719 CysThrLeuGlnProIleLeuLeu---MetGlyGluLeuValLysAspAspValValGlu 737

Qy    1261 AGTTTGTACAAATACCACGATTTAGTTCGTTGGTCATCCTTCGTTCTGCGGCTTGCTGAT 1320
              ::::::            ::::::         |||         ||||||   :::
Db     738 LysValHisTyrProSerAsnMetPheGluLeuValSerLeuSerTrpArgLeuThrAsn 757

Qy    1321 GATTTGGGAACCTCGGTGGAAGAGGTGAGCAGAGGGGATGTGCCGAAATCACTTCAGTGC 1380
           |||      |||         |||   :::||||||               :::   |||
Db     758 AspThrLysThrTyrGlnAlaGluLysAlaArgGlyGlnGlnAlaSerGlyIleAlaCys 777

Qy    1381 TACATGAGTGACTAC---AATGCATCGGAGGCGGAGGCGCGGAAGCACGTGAAATGGCTG 1437
           ||||||   |||         |||:::|||   :::|||   ||||||:::      :::
Db     778 TyrMetLysAspAsnProGlyAlaThrGluGluAspAlaIleLysHisIleCysArgVal 797

Qy    1438 ATAGCGGAGGTGTGGAAGAAGATGAATGCGGAG------AGGGTGTCGAAGGATTCTCCA 1491
           :::               |||:::   :::   |||      :::   |||   |||   |||
Db     798 ValAspArgAlaLeuLysGluAlaSerPheGluTyrPheLysProSerAsnAspIlePro 817

Qy    1492 TTCGGC---AAAGATTTTATAGGATGTGCAGTTGATTTAGGAAGGATGGCGCAGTTGATG 1548
              |||   |||   ||||||                  :::|||      |||:::
Db     818 MetGlyCysLysSerPheIle---------PheAsnLeuArgLeuCysValGlnIlePhe 834

Qy    1549 TACCATAATGGAGATGGGCACGGCACACAACACCCTATTATACATCAACAAATGACCAGA 1608
           |||         ||||||:::|||         :::      |||      :::   :::
Db     835 TyrLysPheIleAspGlyTyrGlyIleAlaAsnGluGluIleLysAspTyrIleArgLys 854

Qy    1609 ACCTTATTCGAGCCC 1623
                 :::|||
Db     855 ValTyrIleAspPro 859

RESULT 12
TASY_TAXCH
ID   TASY_TAXCH     STANDARD;      PRT;   862 AA.
AC   Q9FT37;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Taxadiene synthase (EC 4.2.3.17) (Taxa-4(5),11(12)-diene synthase).
GN   TDC1.
OS   Taxus chinensis (Chinese yew).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Coniferopsida; Coniferales; Taxaceae; Taxus.
OX   NCBI_TaxID=29808;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Callus;
RA   Wang W., Shi Q., Ouyang T., Zhu P., Cheng K.;
RT   "Cloning, expression, and characterization of taxadiene synthase, a
RT   diteroene cyclase from Taxus chinensis.";
RL   Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Catalyzes the cyclization of the ubiquitous isoprenoid
CC       intermediate geranylgeranyl diphosphate to taxa-4,11-diene, the
CC       parent olefin with a taxane skeleton.
CC   -!- CATALYTIC ACTIVITY: Geranylgeranyl diphosphate = taxa-4,11-diene +
CC       diphosphate.
CC   -!- PATHWAY: Taxol biosynthesis; first step.
CC   -!- SIMILARITY: Belongs to the terpene synthase family.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
```

```
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by   non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; AY007207; AAG02257.1; -.
DR   HSSP; Q40577; 5EAU.
DR   InterPro; IPR008930; Terp_cyc_toroid.
DR   InterPro; IPR001906; Terp_synth-like.
DR   InterPro; IPR005630; Terpene_synth_C.
DR   InterPro; IPR008949; Terpenoid_synth.
DR   Pfam; PF01397; Terpene_synth; 1.
DR   Pfam; PF03936; Terpene_synth_C; 1.
KW   Taxol biosynthesis; Lyase.
FT   ACT_SITE     758     758        BY SIMILARITY.
FT   ACT_SITE     835     835        BY SIMILARITY.
FT   ACT_SITE     839     839        BY SIMILARITY.
SQ   SEQUENCE   862 AA;  98069 MW;  3A597BAF722BF679 CRC64;

Alignment Scores:
Pred. No.:              1.32e-38        Length:         862
Score:                  635.00          Matches:        160
Percent Similarity:     50.86%          Conservative:   107
Best Local Similarity:  30.48%          Mismatches:     210
Query Match:            21.84%          Indels:         48
DB:                     1               Gaps:           17

US-09-938-956-6 (1-1632) x TASY_TAXCH (1-862)

Qy        145 GATCAAATTCGACAACTTGAGTTGATCGATGACTTGCAGAGGATGGGGCTGTCCGATCAT 204
Db        351 AspLeuLeuGluArgLeuSerLeuValAspAsnIleGluHisLeuGlyIleGlyArgHis 370

Qy        205 TTCCAAAATGAGTTCAAAGAAATCTTGTCCTCTATATATCTCGACCATCACTATTACAAG 264
Db        371 PheLysGlnGluIleLysValAlaLeuAspTyrValTyr------ArgHisTrp------ 386

Qy        265 AACCCTTTTCCAAAAGAAGAAAGG-------------------------------GAT 291
Db        387 ---------------SerGluArgGlyIleGlyTrpGlyArgAspSerLeuValProAsp 401

Qy        292 CTCTACTCCACATCTCTTGCATTTAGGCTCCTCAGAGAACATGGTTTTCAAGTCGCACAA 351
Db        402 LeuAsnThrThrAlaLeuGlyLeuArgThrLeuArgThrHisGlyTyrAspValSerSer 421

Qy        352 GAGGTATTCGATAGTTTCAAGAACGAGGAGGGTGAGTTCAAAGAAAGCCTTAGCGAC--- 408
Db        422 AspValLeuAsnAsnPheLysAspGluAsnGlyArgPhePheSerSerAlaGlyGlnThr 441

Qy        409 ------GACACCAGAGGATTGTTGCAACTGTATGAAGCTTCCTTTCTGTTGACGGAAGGC 462
Db        442 HisValGluLeuArgSerValValIleLeuPheArgAlaSerAspLeuAlaPheProAsp 461

Qy        463 GAAACCACGCTCGAGTCAGCGAGGGAATTCGCCACCAAATTTTTGGAGGAAAAAGTG--- 519
Db        462 GluGlyAlaMetAspAspAlaArgLysPheAlaGluProTyrLeuArgAspAlaLeuAla 481

Qy        520 AACGAGGGTGGTGTTGATGGCGACCTTTTAACAAGAATCGCATATTCTTTGGACATCCCT 579
Db        482 ThrLysIleSerThrAsnThrLysLeuPheLysGluIleGluTyrValValGluTyrPro 501

Qy        580 CTTCATTGGAGGATTAAAAGGCCAAATGCACCTGTGTGGATCGAA--------------- 624
Db        502 TrpHisMetSerIleProArgSerGluAlaArgSerTyrIleAspSerTyrAspAspAsp 521

Qy        625 ------TGGTATAGGAAG------------AGGCCCGACATG---AATCCAGTAGTGTTG 663
Db        522 TyrValTrpGluArgLysThrLeuTyrArgMetProSerLeuSerAsnSerLysCysLeu 541

Qy        664 GAGCTTGCCATACTCGACTTAAATATTGTTCAAGCACAATTTCAAGAAGAGCTCAAAGAA 723
Db        542 GluLeuAlaLysLeuAspPheAsnIleValGlnSerLeuHisGlnGluGluLeuLysLeu 561

Qy        724 TCCTTCAGGTGGTGGAGAAATACTGGGTTTGTTGAGAAGCTGCCCTTCGCAAGGGATAGA 783
Db        562 LeuThrArgTrpTrpLysGluSerGlyMetAlaAsp---IleAsnPheThrArgHisArg 580

Qy        784 CTGGTGGAATGCTACTTTTGGAATACTGGGATCATCGAGCCACGTCAGCATGCAAGTGCA 843
Db        581 ValAlaGluValTyrPhe---SerSerAlaThrPheGluPro---GluTyrSerAlaThr 598

Qy        844 AGGATAATGATGGGCAAAGTCAACGCTCTGATTACGGTGATCGATGATATTTATGATGTC 903
Db        599 ArgIleAlaPheThrLysIleGlyCysLeuGlnValLeuPheAspAspMetAlaAspIle 618

Qy        904 TATGGCACCTTAGAAGAACTCGAACAATTCACTGACCTCATTCGAAGATGGGATATAAAC 963
Db        619 PheAlaThrLeuAspGluLeuLysSerPheThrGluGlyValLysArgTrpAspThrSer 638

Qy        964 TCAATCGACCAACTTCCCGATTACATGCAACTGTGCTTTCTTGCACTCAACAACTTCGTC 1023
Db        639 LeuLeuHisGluIleProGluCysMetGlnThrCysPheLysValTrpPheLysLeuIle 658

Qy       1024 GATGATACATCGTACGATGTTATGAAGGAGAAAGGCGTCAACGTTATACCCTACCTGCGG 1083
Db        659 GluGluValAsnAsnAspValValLysValGlnGlyArgAspMetLeuAlaHisIleArg 678

Qy       1084 CAATCGTGGGTTGATTTGGCGGATAAGTATATGGTAGAGGCACGGTGGTTCTACGGCGGG 1143
Db        679 LysProTrpGluLeuTyrPheAsnCysTyrValGlnGluArgGluTrpLeuAspAlaGly 698

Qy       1144 CACAAACCAAGTTTGGAAGAGTATTTGGAGAACTCATGGCAGTCGATAAGT---GGGCCC 1200
Db        699 TyrIleProThrPheGluGluTyrLeuLysThrTyrAlaIleSerValGlyLeuGlyPro 718

Qy       1201 TGTATGTTAACGCACATATTCTTCCGAGTAACAGATTCGTTCACAAAGGAGACCGTCGAC 1260
Db        719 CysThrLeuGlnProIleLeuLeu---MetGlyGluLeuValLysAspAspValValGlu 737

Qy       1261 AGTTTGTACAAATACCACGATTTAGTTCGTTGGTCATCCTTCGTTCTGCGGCTTGCTGAT 1320
Db        738 LysValHisTyrProSerAsnMetPheGluLeuValSerLeuSerTrpArgLeuThrAsn 757

Qy       1321 GATTTGGGAACCTCGGTGGAAGAGGTGAGCAGAGGGGATGTGCCGAAATCACTTCAGTGC 1380
Db        758 AspThrLysThrTyrGlnAlaGluLysAlaArgGlyGlnGlnAlaSerGlyIleAlaCys 777

Qy       1381 TACATGAGTGAC---TACAATGCATCGGAGGCGGAGGCGCGGAAGCACGTGAAATGGCTG 1437
Db        778 TyrMetLysAspAsnLeuGlyAlaThrGluGluAspAlaIleLysHisIleCysArgVal 797

Qy       1438 ATAGCGGAGGTGTGGAAGAAGATGAATGCGGAG------AGGGTGTCGAAGGATTCTCCA 1491
Db        798 ValAspArgAlaLeuLysGluAlaSerPheGluTyrPheLysProSerAsnAspIlePro 817

Qy       1492 TTCGGC---AAAGATTTTATAGGATGTGCAGTTGATTTAGGAAGGATGGCGCAGTTGATG 1548
Db        818 MetGlyCysLysSerPheIle---------PheAsnLeuArgLeuCysValGlnIlePhe 834

Qy       1549 TACCATAATGGAGATGGGCACGGCACACAACACCCTATTATACATCAACAAATGACCAGA 1608
Db        835 TyrLysPheIleAspGlyTyrGlyIleAlaAsnGluGluIleLysAspTyrIleArgLys 854

Qy       1609 ACCTTATTCGAGCCC 1623
Db        855 ValTyrIleAspPro 859

RESULT 13
TASY_TAXBA
ID   TASY_TAXBA     STANDARD;      PRT;   862 AA.
AC   Q93YA3;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
```

```
DE   Taxadiene synthase (EC 4.2.3.17) (Taxa-4(5),11(12)-diene synthase).
GN   TDC1 OR TASY.
OS   Taxus baccata (English yew).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Coniferopsida; Coniferales; Taxaceae; Taxus.
OX   NCBI_TaxID=25629;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Goerhardt B.;
RL   Thesis (2001), Technische Universitaet Berlin, Germany.
CC   -!- FUNCTION: Catalyzes the cyclization of the ubiquitous isoprenoid
CC       intermediate geranylgeranyl diphosphate to taxa-4,11-diene, the
CC       parent olefin with a taxane skeleton.
CC   -!- CATALYTIC ACTIVITY: Geranylgeranyl diphosphate = taxa-4,11-diene +
CC       diphosphate.
CC   -!- PATHWAY: Taxol biosynthesis; first step.
CC   -!- SIMILARITY: Belongs to the terpene synthase family.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AJ320538; CAC42773.1; -.
DR   InterPro; IPR008930; Terp_cyc_toroid.
DR   InterPro; IPR001906; Terp_synth-like.
DR   InterPro; IPR005630; Terpene_synth_C.
DR   InterPro; IPR008949; Terpenoid_synth.
DR   Pfam; PF01397; Terpene_synth; 1.
DR   Pfam; PF03936; Terpene_synth_C; 1.
KW   Taxol biosynthesis; Lyase.
FT   ACT_SITE    758    758       BY SIMILARITY.
FT   ACT_SITE    835    835       BY SIMILARITY.
FT   ACT_SITE    839    839       BY SIMILARITY.
SQ   SEQUENCE   862 AA;  98048 MW;  B37835AFFDC1DEFE CRC64;

Alignment Scores:
Pred. No.:               1.43e-38       Length:         862
Score:                   634.50         Matches:        164
Percent Similarity:      50.09%         Conservative:   109
Best Local Similarity:   30.09%         Mismatches:     215
Query Match:             21.82%         Indels:         57
DB:                      1              Gaps:           18

US-09-938-956-6 (1-1632) x TASY_TAXBA (1-862)

Qy        112 ACTTTGGTGAAGATGGAACTGGAGAAA--------------------------GAAACG 144
              ||||||:::            |||:::|||
Db        331 ThrLeuLeuAsnAsnLeuLeuAspLysPheGlyGlyCysValProCysMetTyrSerIle 350

Qy        145 GATCAAATTCGACAACTTGAGTTGATCGATGACTTGCAGAGGATGGGGCTGTCCGATCAT 204
              |||    :::    :::|||    |||:::|||:::::::::    :::|||:::        |||
Db        351 AspLeuLeuGluArgLeuSerLeuValAspAsnIleGluHisLeuGlyIleGlyArgHis 370

Qy        205 TTCCAAAATGAGTTCAAAGAAATCTTGTCCTCTATATATCTCGACCATCACTATTACAAG 264
              |||:::    |||    |||        |||        :::|||          |||:::
Db        371 PheLysGlnGluIleLysValAlaLeuAspTyrValTyr------ArgHisTrp------ 386

Qy        265 AACCCTTTTCCAAAAGAAGAAAGG-------------------------------GAT 291
                                  ||||||                                     |||
Db        387 --------------SerGluArgGlyIleGlyTrpGlyArgAspSerLeuValProAsp 401

Qy        292 CTCTACTCCACATCTCTTGCATTTAGGCTCCTCAGAGAACATGGTTTTCAAGTCGCACAA 351
              |||    :::|||:::|||          |||    ||||||    ||||||:::    |||:::
Db        402 LeuAsnThrThrAlaLeuGlyLeuArgThrLeuArgThrHisGlyTyrAspValSerSer 421

Qy        352 GAGGTATTCGATAGTTTCAAGAACGAGGAGGGTGAGTTCAAAGAAAGCCTTAGCGAC--- 408
              :::|||    ::::::|||||||:::|||    |||    |||          |||
Db        422 AspValLeuAsnAsnPheLysAspGluAsnGlyArgPhePheSerSerAlaGlyGlnThr 441

Qy        409 ------GACACCAGAGGATTGTTGCAACTGTATGAAGCTTCCTTTCTGTTGACGGAAGGC 462
                    :::    |||    ::::::    |||:::    ||||||    |||
Db        442 HisValGluLeuArgSerValValAsnLeuPheArgAlaSerAspLeuAlaPheProAsp 461

Qy        463 GAAACCACGCTCGAGTCAGCGAGGGAATTCGCCACCAAATTTTTGGAGGAAAAAGTG--- 519
              |||        ::::::    ||||||:::|||||          :::|||    :::    :::
Db        462 GluGlyAlaMetAspAspAlaArgLysPheAlaGluProTyrLeuArgAspAlaLeuAla 481

Qy        520 AACGAGGGTGGTGTTGATGGCGACCTTTTAACAAGAATCGCATATTCTTTGGACATCCCT 579
                 :::            :::          |||          |||    |||    ::::::    |||
Db        482 ThrLysIleSerThrAsnThrLysLeuTyrLysGluIleGluTyrValValGluTyrPro 501

Qy        580 CTTCATTGGAGGATTAAAAGGCCAAATGCACCTGTGTGGATCGAA--------------- 624
                 |||      |||    |||          |||          :::|||:::
Db        502 TrpHisMetSerIleProArgLeuGluAlaArgSerTyrIleAspSerTyrAspAspAsp 521

Qy        625 ------TGGTATAGGAAG-----------AGGCCCGACATG---AATCCAGTAGTGTTG 663
                    |||    ||||||                   |||    :::    |||          |||
Db        522 TyrValTrpGlnArgLysThrLeuTyrArgMetProSerLeuSerAsnSerLysCysLeu 541

Qy        664 GAGCTTGCCATACTCGACTTAAATATTGTTCAAGCACAATTTCAAGAAGAGCTCAAAGAA 723
              |||||||||    ||||||    ||||||||||||:::          ||||||||||||||
Db        542 GluLeuAlaLysLeuAspPheAsnIleValGlnSerLeuHisGlnGluGluLeuLysLeu 561

Qy        724 TCCTTCAGGTGGTGGAGAAATACTGGGTTTGTTGAGAAGCTGCCCTTCGCAAGGGATAGA 783
                 |||||||||:::    :::|||         :::    :::    |||    |||    |||
Db        562 LeuThrArgTrpTrpLysGluSerGlyMetAlaAsp---IleAsnPheThrArgHisArg 580

Qy        784 CTGGTGGAATGCTACTTTTGGAATACTGGGATCATCGAGCCACGTCAGCATGCAAGTGCA 843
              :::    |||    ||||||    ::::::          ||||||    :::::::::::
Db        581 ValAlaGluValTyrPhe---SerSerAlaThrPheGluPro---GluTyrSerAlaThr 598

Qy        844 AGGATAATGATGGGCAAAGTCAACGCTCTGATTACGGTGATCGATGATATTTATGATGTC 903
              ||||||          |||:::          |||          :::    ||||||:::    |||:::
Db        599 ArgIleAlaPheThrLysIleGlyCysLeuGlnValLeuPheAspAspMetAlaAspIle 618

Qy        904 TATGGCACCTTAGAAGAACTCGAACAATTCACTGACCTCATTCGAAGATGGGATATAAAC 963
              :::    ||||||:::||||||:::    ||||||:::    ::::::|||||||||    :::
Db        619 PheAlaThrLeuAspGluLeuLysSerPheThrGluGlyValLysArgTrpAspThrSer 638

Qy        964 TCAATCGACCAACTTCCCGATTACATGCAACTGTGCTTTCTTGCACTCAACAACTTCGTC 1023
                 :::    ::::::|||:::    ||||||    ||||||                      :::
Db        639 LeuLeuHisGluIleProGluCysMetGlnThrCysPheLysValTrpPheLysLeuMet 658

Qy       1024 GATGATACATCGTACGATGTTATGAAGGAGAAAGGCGTCAACGTTATACCCTACCTGCGG 1083
              ::::::    :::    ||||||:::|||    :::|||    :::::::::    ::::::|||
Db        659 GluGluValAsnAsnAspValValLysValGlnGlyArgAspMetLeuAlaHisIleArg 678

Qy       1084 CAATCGTGGGTTGATTTGGCGGATAAGTATATGGTAGAGGCACGGTGGTTCTACGGCGGG 1143
              :::    |||          :::    |||:::    |||          |||          |||
Db        679 LysProTrpGluLeuTyrPheAsnCysTyrValGlnGluArgGluTrpLeuGluAlaGly 698

Qy       1144 CACAAACCAAGTTTGGAAGAGTATTTGGAGAACTCATGGCAGTCGATAAGT---GGGCCC 1200
              :::    |||:::    ||||||||||||:::          |||:::          ||||||
Db        699 TyrIleProThrPheGluGluTyrLeuLysThrTyrAlaIleSerValGlyLeuGlyPro 718

Qy       1201 TGTATGTTAACGCACATATTCTTCCGAGTAACAGATTCGTTCACAAAGGAGACCGTCGAC 1260
              |||    |||          |||          :::    :::              :::    |||:::
Db        719 CysThrLeuGlnProIleLeuLeu---MetGlyGluLeuValLysAspAspValValGlu 737

Qy       1261 AGTTTGTACAAATACCACGATTTAGTTCGTTGGTCATCCTTCGTTCTGCGGCTTGCTGAT 1320
              ::::::          ::::::          |||          |||    ||||||    :::
Db        738 LysValHisTyrProSerAsnMetPheGluLeuValSerLeuSerTrpArgLeuThrAsn 757

Qy       1321 GATTTGGGAACCTCGGTGGAAGAGGTGAGCAGAGGGGATGTGCCGAAATCACTTCAGTGC 1380
              |||          |||          |||    :::||||||               :::    |||
Db        758 AspThrLysThrTyrGlnAlaGluLysAlaArgGlyGlnGlnAlaSerGlyIleAlaCys 777

Qy       1381 TACATGAGTGACTAC---AATGCATCGGAGGCGGAGGCGCGGAAGCACGTGAAATGGCTG 1437
              ||||||    |||          |||:::|||    :::|||    ||||||:::          :::
Db        778 TyrMetLysAspAsnProGlyAlaThrGluGluAspAlaIleLysHisIleCysArgVal 797

Qy       1438 ATAGCGGAGGTGTGGAAGAAGATGAATGCGGAG------AGGGTGTCGAAGGATTCTCCA 1491
```

```
                :::               |||:::  :::   |||      :::   |||   |||   |||
Db        798 ValAspArgAlaLeuLysGluAlaSerPheGluTyrPheLysProSerAsnAspIlePro 817

Qy       1492 TTCGGC---AAAGATTTTATAGGATGTGCAGTTGATTTAGGAAGGATGGCGCAGTTGATG 1548
                  |||   |||   ||||||           :::|||           |||:::
Db        818 MetGlyCysLysSerPheIle---------PheAsnLeuArgLeuCysValGlnIlePhe 834

Qy       1549 TACCATAATGGAGATGGGCACGGCACACAACACCCTATTATACATCAACAAATGACCAGA 1608
                |||         ||||||:::|||       :::      |||          :::   :::
Db        835 TyrLysPheIleAspGlyTyrGlyIleAlaAsnGluGluIleLysAspTyrIleArgLys 854

Qy       1609 ACCTTATTCGAGCCC 1623
                       :::|||
Db        855 ValTyrIleAspPro 859

RESULT 14
YK82_YEAST
ID   YK82_YEAST     STANDARD;      PRT;  1169 AA.
AC   P36170;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Hypothetical 122.2 kDa protein in SIR1 3'region precursor.
GN   YKR102W.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Gaillon L., Dujon B.;
RL   Submitted (MAR-1994) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: Cell wall (By similarity).
CC   -!- PTM: Extensively O-glycosylated (By similarity).
CC   -!- SIMILARITY: STRONG, TO YEAST PROTEIN FLO1.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; Z28327; CAA82182.1; -.
DR   PIR; S38181; S38181.
DR   GermOnline; 140081; -.
DR   SGD; S0001810; FLO10.
KW   Hypothetical protein; Glycoprotein; Repeat; Cell wall; Signal.
FT   SIGNAL        1     26       POTENTIAL.
FT   CHAIN        27   1169       HYPOTHETICAL PROTEIN YKR102W.
FT   DOMAIN      274    609       SER/THR-RICH.
FT   CARBOHYD    122    122       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    157    157       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    279    279       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    389    389       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    452    452       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    515    515       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    578    578       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    656    656       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    686    686       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    879    879       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1092   1092       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1099   1099       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   1169 AA;  122164 MW;  940C6DFB3569C669 CRC64;

Alignment Scores:
Pred. No.:               0.0232       Length:        1169
Score:                   138.50       Matches:       105
Percent Similarity:      39.29%       Conservative:  73
Best Local Similarity:   23.18%       Mismatches:    164
Query Match:             4.77%        Indels:        112
DB:                      1            Gaps:          21

US-09-938-956-6 (1-1632) x YK82_YEAST (1-1169)

Qy       1558 CATTATGGTACATCAACTGCG---------------CCATCCTTCCTAAATCAACTGCAC 1514
                :::      :::|||||||||                    ||||||   :::
Db        569 TyrValSerSerSerThrAlaAlaAlaAsnTyrThrSerSerPheSerSerSerSerGlu 588

Qy       1513 ATCCTATAAAATCTTTGCCGAATGGAGAATCCTTCGACACCCTCTCCGCATTCATCTTCT 1454
                :::                     |||:::   ||||||   :::|||:::::::::|||
Db        589 ValCysThrGluCysThrGluThrGluSerThrSerThrSerThrProTyrAlaThrSer 608

Qy       1453 TCCACACC-TCCGCTATCAGCCATTTCACGTGCTTCCGCGCCTCCGCCTCCGATGCATTG 1395
                ||||||   :::|||               |||   |||   ||||||   |||:::   :::
Db        609 SerThrGlyThrAla------------ThrSerPheThrAlaSerThrSerAsnThrMet 624

Qy       1394 TAGTCACTCATGTAGCACTGAAGT------GATTTCGGCACATCCCCTCTGCTCACCTCT 1341
                   ||||||:::        :::          |||      |||      ::::::
Db        625 ThrSerLeuValGlnThrAspThrThrValSerPheSerLeuSerSerThrValSerGlu 644

Qy       1340 TCCACCGAGGTTCCCAAATCATCAGCAAGCCGCAGA--------------ACGAAGGAT 1296
                   |||      |||   ||||||                          :::   :::
Db        645 HisThrAsnAlaProThrSerSerValGluSerAsnAlaSerThrPheIleSerSerAsn 664

Qy       1295 GACCAACGAACTAAATCGTGGTATTTGTACAAACTGTCGACGGTCTCCTTTGTG------ 1242
                         ||||||:::              :::   :::::::::   :::
Db        665 LysGlySerValLysSerTyrValThrSerSerIleHisSerIleThrProMetTyrPro 684

Qy       1241 ---AACGAATCTGTTACTCGGAAGAATATGTGCGTTAACATACAGGGCCCACTTATCGAC 1185
                   |||::::::||||||         ::::::             :::      |||:::
Db        685 SerAsnGlnThrValThrSerSerSerVal---------ValSerThrProIleThrSer 701

Qy       1184 TGCCATGAGTTCTCCAAATACTCTTCCAAACTTGGTTTGTGCCCGCCGTAGAACCACCGT 1125
                      |||   |||:::   |||:::   :::   :::   |||
Db        702 ------GluSerSerGluSerSerAlaSerValThrIleLeuPro--------------- 714

Qy       1124 GCCTCTACCATATACTTATCCGCCAAA---TCAACCCACGATTGCCGCAGGTAGGGTATA 1068
                   |||||||||               |||   ||||||         :::         |||
Db        715 ---SerThrIleThrSerGluPheLysProSerThrMetLysThrLysValValSerIle 733

Qy       1067 ACGTTGACGCCTTTCTCCTTCATAACATCGTACGATGTATCATCGACGAAGTTGTTGAGT 1008
                :::   :::|||   :::   |||||||||||||||   :::|||
Db        734 SerSerSerProThrAsnLeuIleThrSerTyrAspThrThrSer--------------- 748

Qy       1007 GCAAGAAAGCACAGTTGCATGTAATCGGGAAGTTGGTCGATTGAGTTTATATCCCATCTT 948
                      |||   |||   :::   |||   :::   |||:::      ||||||   :::
Db        749 ------LysAspSerThrValGlySerSerThrSerSerValSerLeuIleSerSerIle 766

Qy        947 CGAATGAGGTCAGTGAATTGTTCGAGTTCTTCTAAGGTGCCATAGACATCATAAATATCA 888
                   :::   |||      :::||||||   ::::::      :::|||   :::|||
Db        767 SerLeuProSerSerTyrSerAlaSerSerGluGlnIlePheHisSerSerIleValSer 786

Qy        887 TCGATCACCGTAATCAGAGCGTTGACTTTGCCCATCATTATCCTTGCACTTGCATGCTGA 828
                |||
Db        787 Ser--------------------------------------------------------- 787

Qy        827 CGTGGCTCGATGATCCCAGTATTCCAAAAGTAGCATTCCACCAGTCTATCCCTTGCGAAG 768
                   |||      :::         |||               ||||||   :::|||   ::::::
Db        788 AsnGlyGlnAlaLeuThrSerPheSer---------SerThrLysValSerSerSerGlu 804

Qy        767 GGCAGCTTCTCAACAAACCCAGTATTTCTCCACCACCTGAAGGATTCTTTGAGCTCTTCT 708
                   |||   |||                  |||         :::   ||||||
Db        805 SerSerGluSer------------------HisArgThrSerProThrThrSerSerGlu 818

Qy        707 TGAAATTGTGCTTGAACAATATTTAAGTCGAGTATGGCAAGCTCCAACACTACTGGATTC 648
                               :::      |||||||||            :::|||
Db        819 ---------------SerGlyIleLysSerSerGlyValGluIleGluSerThrSerThr 833

Qy        647 ATGTCGGGCCTCTTCCTATACCATTCGATCCACACAGGTGCATTTGGCCTTTTAATCCTC 588
                |||      |||   :::|||               |||      :::      |||
Db        834 SerSer------PheSerPheHisGluThrSerThrAlaSerThrSerValGlnIle--- 850

Qy        587 CAATGAAGAGGGATGTCCAAAGAATATGCGATTCTTGTTAAAAGGTCGCCATCAACACCA 528
                            |||   ::::::               :::||||||:::|||
```

```
Db        851 ---------------SerSerGlnPheVal---------------ThrProSerSerPro 860

Qy        527 CCCTCGTTC--------------------------------------------------- 519
                   |||
Db        861 IleSerThrValAlaProArgSerThrGlyLeuAsnSerGlnThrGluSerThrAsnSer 880

Qy        518 ---------ACTTTTTCCTCCAAAAATTTGGTGGCGAATTCCCTCGCTGACTCGAGCGTG 468
                       |||  ||||||:::|||      |||:::   :::       ||||||
Db        881 SerLysGluThrMetSerSerGluAsn---SerAlaSerValMetProSerSerSerAla 899

Qy        467 GTTTCGCCT---------TCCGTCAACAGAAAGGAAGCTTCATACAGTTGCAACAATCCT 417
                 ||||||         |||          |||  |||          :::
Db        900 ThrSerProLysThrGlyLysValThrSerAspGluThrSerSerGlyPheSerArgAsp 919

Qy        416 CTGGTGTCGTCGCTAAGGCTTTCTTTGAACTCACCCTCCTCGTTCTTGAAACTATCGAAT 357
                    :::      |||::::::       :::||||||:::
Db        920 ArgThrThrValTyrArgMetThrSerGluThrProSerThr---------AsnGluGln 936

Qy        356 ACCTCTTGTGCGACTTGAAAACCATGTTCTCTGAGGAGC 318
                 |||:::      |||        |||        |||
Db        937 ThrThrLeuIleThrValSerSerCysGluSerAsnSer 949
```

RESULT 15

```
RBP1_PLAVB
ID   RBP1_PLAVB     STANDARD;      PRT;  2869 AA.
AC   Q00798;
DT   01-APR-1993 (Rel. 25, Created)
DT   01-APR-1993 (Rel. 25, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Reticulocyte binding protein 1 precursor.
GN   RBP1.
OS   Plasmodium vivax (strain Belem).
OC   Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX   NCBI_TaxID=31273;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92315338; PubMed=1617731;
RA   Galinski M.R., Medina C.C., Ingravallo P., Barnwell J.W.;
RT   "A reticulocyte-binding protein complex of Plasmodium vivax
RT   merozoites.";
RL   Cell 69:1213-1226(1992).
CC   -!- FUNCTION: Involved in reticulocyte adhesion. Specifically binds to
CC       human reticulocyte cells.
CC   -!- SUBUNIT: Homodimer (Potential).
CC   -!- SUBCELLULAR LOCATION: Membrane-bound.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; M88097; AAA29743.1; -.
KW   Malaria; Receptor; Signal; Transmembrane.
FT   SIGNAL          1     17       POTENTIAL.
FT   CHAIN          18   2869       RETICULOCYTE BINDING PROTEIN 1.
FT   DOMAIN         18   2807       EXTRACELLULAR.
FT   TRANSMEM     2808   2826       POTENTIAL.
FT   DOMAIN       2827   2869       CYTOPLASMIC.
FT   SITE         1030   1032       CELL ATTACHMENT SITE (POTENTIAL).
FT   SITE         2599   2601       CELL ATTACHMENT SITE (POTENTIAL).
SQ   SEQUENCE   2869 AA;  330213 MW;  B9DBE442205EBCFF CRC64;

Alignment Scores:
Pred. No.:               0.108        Length:          2869
Score:                   130.50       Matches:         111
Percent Similarity:      35.55%       Conservative:    76
Best Local Similarity:   21.10%       Mismatches:      176
Query Match:             4.49%        Indels:          163
DB:                      1            Gaps:            25

US-09-938-956-6 (1-1632) x RBP1_PLAVB (1-2869)

Qy         49 ATCCAATCGCTTCTCAGTGACTATAAGGAGGACAAACACGTGATTAGGGCTTCTGAGCTG 108
              :::|||::::::            ||||||||         ::::::   :::   |||
Db        630 LeuGlnThrValGluLysPheTyrLysGlu---------IleLeuAspSerLysGluLys 646

Qy        109 GTCACTTTGGTGAAGATGGAACTGGAGAAAGAAACGGATCAAATTCGACAACTTGAGTTG 168
              :::       :::|||:::|||   |||||         :::|||    :::|||:::
Db        647 IleTyrGluLeuLysIleGluPheGluLysSerValThrGluIleAsnArgLeuGln--- 665

Qy        169 ATCGATGACTTGCAGAGGATGGGGCTGTCCGATCATTTCCAAAATGAGTTCAAAGAAATC 228
                                                |||    :::   :::   ||||||||
Db        666 ------------AspGlyGluSerAlaArgAspLeuHisGluGluGlnIleLysGluIle 681

Qy        229 TTGTCCTCTATATATCTCGACCATCACTATTACAAGAACCCTTTTCCA---AAAGAAGAA 285
              |||      :::            ||||||   |||                |||   :::
Db        682 LeuAspLysMetAlaLysLysValHisTyrLeuLysGluLeuLeuSerLeuLysGlyLys 701

Qy        286 AGGGATCTCTACTCCACATCTCTTGCATTTAGGCTCCTCAGAGAA--------------- 330
                    :::|||   |||    :::            ||||||
Db        702 SerSerValTyrPheThrGluMetAsn---GluLeuLeuAsnThrAlaSerTyrAspAsn 720

Qy        331 ---CATGGTTTTCAAGTCGCACAAGAGGTATTCGATAGTTTCAAGAACGAGGAGGGTGAG 387
                    ||||||           :::|||   |||:::         |||
Db        721 MetGluGlyPheSerAlaLysLysGluLysAlaAspAsnAspIleAsn------AlaLeu 738

Qy        388 TTCAAAGAAAGCCTTAGCGACGACACCAGAGGATTGTTGCAACTGTATGAAGCTTCCTTT 447
              :::               :::|||         |||::::::         |||      |||
Db        739 TyrAsnSerValTyrArgGluAspIleAsnAlaLeuIleGluGluValGlu---LysPhe 757

Qy        448 CTGTTGACGGAAGGCGAAACCACGCTCGAGTCAGCGAGGGAATTCGCCACCAAATTTTTG 507
              :::               |||:::|||||||||      ::::::         :::   :::
Db        758 ValThrGluAsnLysGluSerThrLeuGluMetLeuLysAsp---------GluGluMet 774

Qy        508 GAGGAAAAAGTGAACGAGGGTGGT-----------------------GTTGATGGCGAC 543
              |||||||||:::   :::                               |||      |||
Db        775 GluGluLysLeuGlnAspAlaLysGluThrPheAlaLysLeuAsnPheValSerAspAsp 794

Qy        544 CTTTTAACAAGAATCGCATATTCTTTGGACATCCCTCTTCATTGGAGGATTAAAAGGCCA 603
                 ||||||   :::
Db        795 LysLeuThrAspVal--------------------------------------------- 799

Qy        604 AATGCACCTGTGTGGATCGAATGGTATAGGAAGAGGCCCGACATGAATCCAGTAGTGTTG 663
                                     |||   |||         |||:::      |||
Db        800 -----------------------TyrThrLys---------MetSerAlaGluValThr 808

Qy        664 GAGCTTGCCATACTCGACTTAAATATTGTTCAAGCACAATTTCAA--------------- 708
                       :::         |||   |||   ||||||:::
Db        809 AsnAlaGluGlyIleLysLysGluIleAlaGlnLysGlnPheGluAsnValHisLysLys 828

Qy        709 -----------------------------------GAAGAGCTCAAAGAATCCTTCAGG 732
                                                   |||   |||:::   |||   :::
Db        829 MetLysGluPheSerAspAlaPheSerThrLysPheGluAlaLeuGlnAsnSerMetGln 848

Qy        733 TGGTGGAGAAATACTGGG---TTTGTTGAGAAGCTGCCCTTCGCAAGGGATAGACTGGTG 789
                 :::         |||      :::||||||            |||
Db        849 GlnTyrAsnGlnGluGlyAspAlaIleGluLysHisLysGlnAsnArgSerGluLysGlu 868

Qy        790 GAATGCTACTTTTGGAATACTGGGATCATCGAGCCA------------------------ 825
              |||   ||||||   |||       :::   |||
Db        869 GluGluTyrPheLysAsnGluSerValGluGluAspLeuSerArgGluGluThrGluGlu 888

Qy        826 ------------------CGTCAGCATGCAAGTGCAAGGATAATGATGGGCAAAGTCAAC 867
                                :::   :::   |||   |||            |||:::::::::
Db        889 GlnGluTyrThrLysHisLysAsnAsnPheSerArgArg------LysGlyGluIleSer 906

Qy        868 GCTCTGATTACGGTGATCGATGATATTTATGATGTCTATGGCACCTTAGAAGAACTCGAA 927
              |||   ||||||   :::   ::::::
Db        907 AlaGluIleThrAsnMetArgGluVal--------------------------------- 915

Qy        928 CAATTCACTGACCTCATTCGAAGATGGGATATAAACTCAATCGACCAACTTCCCGATTAC 987
                                            ||||||   |||:::         :::|||
Db        916 -----------------------------IleAsnLysIleGluSerGlnLeuAsnTyr 925
```

```
Qy         988 ---------ATGCAACTGTGCTTTCTTGCACTCAACAACTTCGTCGATGATACATCGTAC 1038
                        ::::::      |||      :::   :::      :::   ::::::
Db         926 TyrGlyValIleGluLysTyrPheSerLeuIleGlyAspGlnAsnGluValSerThrAla 945

Qy        1039 GATGTTATGAAGGAGAAAGGCGTCAACGTTATACCCTACCTGCGGCAATCGTGGGTTGAT 1098
                     :::|||||||||   |||:::                     |||      |||
Db         946 LysAlaLeuLysGluLysIleValSer------------------AspSerLeuArgAsp 959

Qy        1099 TTGGCGGATAAGTATATGGTAGAGGCACGGTGGTTCTACGGCGGGCACAAACCAAGTTTG 1158
                     |||:::|||      |||   :::                  |||   |||
Db         960 LysIleAspGlnTyrGluThrGluPheLys---------------GluLysThrSerAla 974

Qy        1159 GAAGAGTATTTGGAGAACTCATGGCAGTCGATAAGTGGGCCCTGTATGTTAACGCACATA 1218
                  |||         ::::::   ||||||:::|||
Db         975 ValGluAsnThrValSerThrIleGlnSerLeuSer------------------------ 986

Qy        1219 TTCTTCCGAGTAACAGATTCGTTCACAAAG------GAGACCGTCGACAGTTTGTACAAA 1272
                     :::      ||||||      :::         :::::::::::         |||
Db         987 ------LysAlaIleAspSerLeuLysArgLeuAsnGlySerIleAsnAsnCysLysLys 1004

Qy        1273 TACCAC---------GATTTAGTTCGTTGGTCATCCTTCGTTCTGCGGCTTGCTGATGAT 1323
               |||:::         ||||||:::|||               ||||||
Db        1005 TyrAsnThrAspIleAspLeuLeuArgSerLysIleLysThrLeuArg------------ 1020

Qy        1324 TTGGGAACCTCGGTGGAAGAGGTG---------------AGCAGAGGGGAT--------- 1359
                              |||||||||                  |||||||||
Db        1021 ---------------GluGluValGlnLysGluMetProLysArgGlyAspLysCysGly 1035

Qy        1360 ------------------GTGCCGAAATCACTTCAGTGCTACATGAGTGACTACAATGCA 1401
                                 :::   |||||||||:::            |||         |||
Db        1036 GluAsnThrThrAlaLeuLeuLeuLysSerLeuArgAspLysMetGlyLysIleAsnGlu 1055

Qy        1402 TCGGAGGCGGAGGCGCGG 1419
                        :::   |||
Db        1056 LysLeuAsnAspGlyArg 1061
```

Search completed: February 25, 2004, 05:43:28
Job time : 75.5 secs

```
OM nucleic - protein search, using frame_plus_n2p model

Run on:          February 25, 2004, 05:37:34 ; Search time 79.5 Seconds
                                         (without alignments)
                                         11600.441 Million cell updates/sec

Title:           US-09-938-956-6
Perfect score:   2908
Sequence:        1 atgagacgatccggaaacta..........tattcgagccctttgcatga 1632

Scoring table:   BLOSUM62
                 Xgapop 10.0 , Xgapext  0.5
                 Ygapop 10.0 , Ygapext  0.5
                 Fgapop  6.0 , Fgapext  7.0
                 Delop   6.0 , Delext   7.0

Searched:        1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters:      3172214

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US09938956/runat_24022004_144659_3047/app_query.fasta_1.1799
-DB=A_Geneseq_29Jan04 -QFMT=fastan -SUFFIX=rag -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09938956_@CGN_1_1_154_@runat_24022004_144659_3047 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :       A_Geneseq_29Jan04:*
                 1:  geneseqp1980s:*
                 2:  geneseqp1990s:*
                 3:  geneseqp2000s:*
                 4:  geneseqp2001s:*
                 5:  geneseqp2002s:*
                 6:  geneseqp2003as:*
                 7:  geneseqp2003bs:*
                 8:  geneseqp2004s:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2881	99.1	543	5	ABB75749	Abb75749 Mentha sp
2	2876	98.9	599	2	AAR74291	Aar74291 Spearmint
3	2876	98.9	599	3	AAY90838	Aay90838 Spearmint
4	2876	98.9	599	3	AAY70716	Aay70716 Mint 4S-1
5	2876	98.9	599	6	AAE29781	Aae29781 M. spicat
6	2825	97.1	599	3	AAB29400	Aab29400 Spearmint
7	1636.5	56.3	591	2	AAY02332	Aay02332 Computer
8	1634.5	56.2	591	2	AAY02335	Aay02335 Computer
9	1633.5	56.2	591	2	AAY02309	Aay02309 A sage 1,
10	1633.5	56.2	591	2	AAY02331	Aay02331 Computer
11	1633.5	56.2	591	3	AAY90839	Aay90839 Sage 1,8
12	1632.5	56.1	591	2	AAY02336	Aay02336 Computer
13	1632.5	56.1	591	2	AAY02329	Aay02329 Computer
14	1631.5	56.1	591	2	AAY02330	Aay02330 Computer
15	1631.5	56.1	591	2	AAY02337	Aay02337 Computer
16	1631.5	56.1	591	2	AAY02333	Aay02333 Computer
17	1630.5	56.1	591	2	AAY02338	Aay02338 Computer
18	1630.5	56.1	591	2	AAY02334	Aay02334 Computer
19	1553.5	53.4	589	2	AAY02310	Aay02310 A sage (+
20	1553.5	53.4	590	3	AAY90857	Aay90857 Sage +sab
21	1553.5	53.4	590	6	AAE29782	Aae29782 S. offici
22	1552.5	53.4	590	2	AAY02321	Aay02321 Computer
23	1552.5	53.4	590	2	AAY02322	Aay02322 Computer
24	1551.5	53.4	590	2	AAY02325	Aay02325 Computer
25	1551.5	53.4	590	2	AAY02323	Aay02323 Computer
26	1550.5	53.3	590	2	AAY02326	Aay02326 Computer
27	1550.5	53.3	590	2	AAY02328	Aay02328 Computer
28	1550.5	53.3	590	2	AAY02324	Aay02324 Computer
29	1550.5	53.3	590	2	AAY02327	Aay02327 Computer
30	1547.5	53.2	589	2	AAY02320	Aay02320 Computer
31	1471	50.6	598	2	AAY02314	Aay02314 Computer
32	1469	50.5	598	2	AAY02317	Aay02317 Computer
33	1469	50.5	598	2	AAY02318	Aay02318 Computer
34	1468	50.5	598	2	AAY02308	Aay02308 A (+)-bor
35	1468	50.5	598	3	AAY90840	Aay90840 Sage born
36	1467	50.4	598	2	AAY02316	Aay02316 Computer
37	1467	50.4	598	2	AAY02311	Aay02311 Computer
38	1466	50.4	598	2	AAY02312	Aay02312 Computer
39	1466	50.4	598	2	AAY02319	Aay02319 Computer
40	1466	50.4	598	2	AAY02313	Aay02313 Computer
41	1465	50.4	598	2	AAY02315	Aay02315 Computer
42	1203.5	41.4	583	5	AAM50097	Aam50097 Armopha-4
43	1161.5	39.9	562	3	AAG51896	Aag51896 Arabidops
44	1161.5	39.9	566	3	AAG51895	Aag51895 Arabidops
45	1161.5	39.9	598	3	AAG51894	Aag51894 Arabidops

ALIGNMENTS

```
RESULT 1
ABB75749
ID   ABB75749 standard; protein; 543 AA.
XX
AC   ABB75749;
XX
DT   24-JUN-2002  (first entry)
XX
DE   Mentha spicata limonene synthase.
XX
KW   Limonene synthase; enzyme; monoterpene; cyclic terpenoid; flavour;
KW   fragrance; plant.
XX
OS   Mentha spicata.
XX
PN   WO200220815-A2.
XX
PD   14-MAR-2002.
XX
PF   28-AUG-2001; 2001WO-US026853.
XX
PR   01-SEP-2000; 2000US-0229858P.
PR   01-SEP-2000; 2000US-0229907P.
XX
PA   (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
PI   Dicosimo DJ,  Koffas M,  Wang S;
XX
DR   WPI; 2002-339805/37.
DR   N-PSDB; ABL53856.
XX
PT   Producing cyclic terpenoids, e.g. monoterpenes, which are useful in the
PT   fragrance or pharmaceutical industry, by employing a transformed
PT   methanotrophic bacterium that metabolize single carbon substrates as a
```

```
PT   sole carbon source.
XX
PS   Claim 15; Page 61-63; 63pp; English.
XX
CC   The present sequence is a Mentha spicata limonene synthase truncated
CC   polypeptide from which 57 N-terminal amino acid residues have been
CC   deleted. Limonene synthase catalyses the conversion of geranyl
CC   pyrophosphate to (-)-limonene. A method of producing high-value
CC   monoterpenes from inexpensive feedstocks uses a C1 metabolising
CC   microorganism that has been transformed with a gene encoding a cyclic
CC   terpene synthase, such as the present limonene synthase. The C1
CC   metabolising microorganism may be a methylotroph or methanotroph, such as
CC   Methylomonas sp. 16a (ATCC PTA 2402), a novel isolate that uses methane
CC   as a sole source of carbon and energy. In an example, Methylomonas sp.
CC   16a transconjugants carrying a gene encoding the truncated limonene
CC   synthase produced detectable amounts (0.5 ppm) of limonene when grown
CC   with methane. Monoterpenes are used in the flavour, fragrance and
CC   pharmaceutical industries
XX
SQ   Sequence 543 AA;

Alignment Scores:
Pred. No.:              1.13e-266     Length:         543
Score:                  2881.00       Matches:        543
Percent Similarity:     100.00%       Conservative:   0
Best Local Similarity:  100.00%       Mismatches:     0
Query Match:            99.07%        Indels:         0
DB:                     5             Gaps:           0

US-09-938-956-6 (1-1632) x ABB75749 (1-543)

Qy        1 ATGAGACGATCCGGAAACTACAACCCTTCTCGTTGGGATGTCAACTTCATCCAATCGCTT 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1 MetArgArgSerGlyAsnTyrAsnProSerArgTrpAspValAsnPheIleGlnSerLeu 20

Qy       61 CTCAGTGACTATAAGGAGGACAAACACGTGATTAGGGCTTCTGAGCTGGTCACTTTGGTG 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       21 LeuSerAspTyrLysGluAspLysHisValIleArgAlaSerGluLeuValThrLeuVal 40

Qy      121 AAGATGGAACTGGAGAAAGAAACGGATCAAATTCGACAACTTGAGTTGATCGATGACTTG 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       41 LysMetGluLeuGluLysGluThrAspGlnIleArgGlnLeuGluLeuIleAspAspLeu 60

Qy      181 CAGAGGATGGGGCTGTCCGATCATTTCCAAAATGAGTTCAAAGAAATCTTGTCCTCTATA 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       61 GlnArgMetGlyLeuSerAspHisPheGlnAsnGluPheLysGluIleLeuSerSerIle 80

Qy      241 TATCTCGACCATCACTATTACAAGAACCCTTTTCCAAAAGAAGAAAGGGATCTCTACTCC 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       81 TyrLeuAspHisHisTyrTyrLysAsnProPheProLysGluGluArgAspLeuTyrSer 100

Qy      301 ACATCTCTTGCATTTAGGCTCCTCAGAGAACATGGTTTTCAAGTCGCACAAGAGGTATTC 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      101 ThrSerLeuAlaPheArgLeuLeuArgGluHisGlyPheGlnValAlaGlnGluValPhe 120

Qy      361 GATAGTTTCAAGAACGAGGAGGGTGAGTTCAAAGAAAGCCTTAGCGACGACACCAGAGGA 420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      121 AspSerPheLysAsnGluGluGlyGluPheLysGluSerLeuSerAspAspThrArgGly 140

Qy      421 TTGTTGCAACTGTATGAAGCTTCCTTTCTGTTGACGGAAGGCGAAACCACGCTCGAGTCA 480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      141 LeuLeuGlnLeuTyrGluAlaSerPheLeuLeuThrGluGlyGluThrThrLeuGluSer 160

Qy      481 GCGAGGGAATTCGCCACCAAATTTTTGGAGGAAAAAGTGAACGAGGGTGGTGTTGATGGC 540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      161 AlaArgGluPheAlaThrLysPheLeuGluGluLysValAsnGluGlyGlyValAspGly 180

Qy      541 GACCTTTTAACAAGAATCGCATATTCTTTGGACATCCCTCTTCATTGGAGGATTAAAAGG 600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      181 AspLeuLeuThrArgIleAlaTyrSerLeuAspIleProLeuHisTrpArgIleLysArg 200

Qy      601 CCAAATGCACCTGTGTGGATCGAATGGTATAGGAAGAGGCCCGACATGAATCCAGTAGTG 660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      201 ProAsnAlaProValTrpIleGluTrpTyrArgLysArgProAspMetAsnProValVal 220

Qy      661 TTGGAGCTTGCCATACTCGACTTAAATATTGTTCAAGCACAATTTCAAGAAGAGCTCAAA 720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      221 LeuGluLeuAlaIleLeuAspLeuAsnIleValGlnAlaGlnPheGlnGluGluLeuLys 240

Qy      721 GAATCCTTCAGGTGGTGGAGAAATACTGGGTTTGTTGAGAAGCTGCCCTTCGCAAGGGAT 780
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      241 GluSerPheArgTrpTrpArgAsnThrGlyPheValGluLysLeuProPheAlaArgAsp 260

Qy      781 AGACTGGTGGAATGCTACTTTTGGAATACTGGGATCATCGAGCCACGTCAGCATGCAAGT 840
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      261 ArgLeuValGluCysTyrPheTrpAsnThrGlyIleIleGluProArgGlnHisAlaSer 280

Qy      841 GCAAGGATAATGATGGGCAAAGTCAACGCTCTGATTACGGTGATCGATGATATTTATGAT 900
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      281 AlaArgIleMetMetGlyLysValAsnAlaLeuIleThrValIleAspAspIleTyrAsp 300

Qy      901 GTCTATGGCACCTTAGAAGAACTCGAACAATTCACTGACCTCATTCGAAGATGGGATATA 960
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      301 ValTyrGlyThrLeuGluGluLeuGluGlnPheThrAspLeuIleArgArgTrpAspIle 320

Qy      961 AACTCAATCGACCAACTTCCCGATTACATGCAACTGTGCTTTCTTGCACTCAACAACTTC 1020
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      321 AsnSerIleAspGlnLeuProAspTyrMetGlnLeuCysPheLeuAlaLeuAsnAsnPhe 340

Qy     1021 GTCGATGATACATCGTACGATGTTATGAAGGAGAAAGGCGTCAACGTTATACCCTACCTG 1080
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      341 ValAspAspThrSerTyrAspValMetLysGluLysGlyValAsnValIleProTyrLeu 360

Qy     1081 CGGCAATCGTGGGTTGATTTGGCGGATAAGTATATGGTAGAGGCACGGTGGTTCTACGGC 1140
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      361 ArgGlnSerTrpValAspLeuAlaAspLysTyrMetValGluAlaArgTrpPheTyrGly 380

Qy     1141 GGGCACAAACCAAGTTTGGAAGAGTATTTGGAGAACTCATGGCAGTCGATAAGTGGGCCC 1200
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      381 GlyHisLysProSerLeuGluGluTyrLeuGluAsnSerTrpGlnSerIleSerGlyPro 400

Qy     1201 TGTATGTTAACGCACATATTCTTCCGAGTAACAGATTCGTTCACAAAGGAGACCGTCGAC 1260
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      401 CysMetLeuThrHisIlePhePheArgValThrAspSerPheThrLysGluThrValAsp 420

Qy     1261 AGTTTGTACAAATACCACGATTTAGTTCGTTGGTCATCCTTCGTTCTGCGGCTTGCTGAT 1320
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      421 SerLeuTyrLysTyrHisAspLeuValArgTrpSerSerPheValLeuArgLeuAlaAsp 440

Qy     1321 GATTTGGGAACCTCGGTGGAAGAGGTGAGCAGAGGGGATGTGCCGAAATCACTTCAGTGC 1380
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      441 AspLeuGlyThrSerValGluGluValSerArgGlyAspValProLysSerLeuGlnCys 460

Qy     1381 TACATGAGTGACTACAATGCATCGGAGGCGGAGGCGCGGAAGCACGTGAAATGGCTGATA 1440
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      461 TyrMetSerAspTyrAsnAlaSerGluAlaGluAlaArgLysHisValLysTrpLeuIle 480

Qy     1441 GCGGAGGTGTGGAAGAAGATGAATGCGGAGAGGGTGTCGAAGGATTCTCCATTCGGCAAA 1500
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      481 AlaGluValTrpLysLysMetAsnAlaGluArgValSerLysAspSerProPheGlyLys 500

Qy     1501 GATTTTATAGGATGTGCAGTTGATTTAGGAAGGATGGCGCAGTTGATGTACCATAATGGA 1560
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      501 AspPheIleGlyCysAlaValAspLeuGlyArgMetAlaGlnLeuMetTyrHisAsnGly 520

Qy     1561 GATGGGCACGGCACACAACACCCTATTATACATCAACAAATGACCAGAACCTTATTCGAG 1620
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      521 AspGlyHisGlyThrGlnHisProIleIleHisGlnGlnMetThrArgThrLeuPheGlu 540

Qy     1621 CCCTTTGCA 1629
            |||||||||
Db      541 ProPheAla 543

RESULT 2
AAR74291
ID   AAR74291 standard; protein; 599 AA.
```

```
XX
AC   AAR74291;
XX
DT   25-MAR-2003  (revised)
DT   24-OCT-1995  (first entry)
XX
DE   Spearmint limonene synthase.
XX
KW   LS; expression; production; predator defence mechanism; enzyme.
XX
OS   Mentha spicata.
XX
PN   WO9511913-A1.
XX
PD   04-MAY-1995.
XX
PF   27-OCT-1994;   94WO-US012379.
XX
PR   28-OCT-1993;   93US-00145941.
XX
PA   (UNIW ) UNIV WASHINGTON STATE RES FOUND.
XX
PI   Croteau RB,  Colby SM;
XX
DR   WPI; 1995-178818/23.
DR   N-PSDB; AAQ90066.
XX
PT   Nucleotide sequence for spearmint limonene synthase - useful as predator
PT   defence mechanism for plants.
XX
PS   Claim 7; Fig 4; 62pp; English.
XX
CC   The sequence is that of the spearmint limonene synthase protein. The
CC   enzyme can be used in plants as a predator defence mechanism. (Updated on
CC   25-MAR-2003 to correct PN field.)
XX
SQ   Sequence 599 AA;

Alignment Scores:
Pred. No.:              3.56e-266     Length:         599
Score:                  2876.00       Matches:        542
Percent Similarity:     100.00%       Conservative:   0
Best Local Similarity:  100.00%       Mismatches:     0
Query Match:            98.90%        Indels:         0
DB:                     2             Gaps:           0

US-09-938-956-6 (1-1632) x AAR74291 (1-599)

Qy          4 AGACGATCCGGAAACTACAACCCTTCTCGTTGGGATGTCAACTTCATCCAATCGCTTCTC 63
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         58 ArgArgSerGlyAsnTyrAsnProSerArgTrpAspValAsnPheIleGlnSerLeuLeu 77

Qy         64 AGTGACTATAAGGAGGACAAACACGTGATTAGGGCTTCTGAGCTGGTCACTTTGGTGAAG 123
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         78 SerAspTyrLysGluAspLysHisValIleArgAlaSerGluLeuValThrLeuValLys 97

Qy        124 ATGGAACTGGAGAAAGAAACGGATCAAATTCGACAACTTGAGTTGATCGATGACTTGCAG 183
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         98 MetGluLeuGluLysGluThrAspGlnIleArgGlnLeuGluLeuIleAspAspLeuGln 117

Qy        184 AGGATGGGGCTGTCCGATCATTTCCAAAATGAGTTCAAAGAAATCTTGTCCTCTATATAT 243
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        118 ArgMetGlyLeuSerAspHisPheGlnAsnGluPheLysGluIleLeuSerSerIleTyr 137

Qy        244 CTCGACCATCACTATTACAAGAACCCTTTTCCAAAAGAAGAAAGGGATCTCTACTCCACA 303
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        138 LeuAspHisHisTyrTyrLysAsnProPheProLysGluGluArgAspLeuTyrSerThr 157

Qy        304 TCTCTTGCATTTAGGCTCCTCAGAGAACATGGTTTTCAAGTCGCACAAGAGGTATTCGAT 363
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        158 SerLeuAlaPheArgLeuLeuArgGluHisGlyPheGlnValAlaGlnGluValPheAsp 177

Qy        364 AGTTTCAAGAACGAGGAGGGTGAGTTCAAAGAAAGCCTTAGCGACGACACCAGAGGATTG 423
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        178 SerPheLysAsnGluGluGlyGluPheLysGluSerLeuSerAspAspThrArgGlyLeu 197

Qy        424 TTGCAACTGTATGAAGCTTCCTTTCTGTTGACGGAAGGCGAAACCACGCTCGAGTCAGCG 483
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        198 LeuGlnLeuTyrGluAlaSerPheLeuLeuThrGluGlyGluThrThrLeuGluSerAla 217

Qy        484 AGGGAATTCGCCACCAAATTTTTGGAGGAAAAAGTGAACGAGGGTGGTGTTGATGGCGAC 543
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        218 ArgGluPheAlaThrLysPheLeuGluGluLysValAsnGluGlyGlyValAspGlyAsp 237

Qy        544 CTTTTAACAAGAATCGCATATTCTTTGGACATCCCTCTTCATTGGAGGATTAAAAGGCCA 603
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        238 LeuLeuThrArgIleAlaTyrSerLeuAspIleProLeuHisTrpArgIleLysArgPro 257

Qy        604 AATGCACCTGTGTGGATCGAATGGTATAGGAAGAGGCCCGACATGAATCCAGTAGTGTTG 663
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        258 AsnAlaProValTrpIleGluTrpTyrArgLysArgProAspMetAsnProValValLeu 277

Qy        664 GAGCTTGCCATACTCGACTTAAATATTGTTCAAGCACAATTTCAAGAAGAGCTCAAAGAA 723
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        278 GluLeuAlaIleLeuAspLeuAsnIleValGlnAlaGlnPheGlnGluGluLeuLysGlu 297

Qy        724 TCCTTCAGGTGGTGGAGAAATACTGGGTTTGTTGAGAAGCTGCCCTTCGCAAGGGATAGA 783
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        298 SerPheArgTrpTrpArgAsnThrGlyPheValGluLysLeuProPheAlaArgAspArg 317

Qy        784 CTGGTGGAATGCTACTTTTGGAATACTGGGATCATCGAGCCACGTCAGCATGCAAGTGCA 843
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        318 LeuValGluCysTyrPheTrpAsnThrGlyIleIleGluProArgGlnHisAlaSerAla 337

Qy        844 AGGATAATGATGGGCAAAGTCAACGCTCTGATTACGGTGATCGATGATATTTATGATGTC 903
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        338 ArgIleMetMetGlyLysValAsnAlaLeuIleThrValIleAspAspIleTyrAspVal 357

Qy        904 TATGGCACCTTAGAAGAACTCGAACAATTCACTGACCTCATTCGAAGATGGGATATAAAC 963
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        358 TyrGlyThrLeuGluGluLeuGluGlnPheThrAspLeuIleArgArgTrpAspIleAsn 377

Qy        964 TCAATCGACCAACTTCCCGATTACATGCAACTGTGCTTTCTTGCACTCAACAACTTCGTC 1023
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        378 SerIleAspGlnLeuProAspTyrMetGlnLeuCysPheLeuAlaLeuAsnAsnPheVal 397

Qy       1024 GATGATACATCGTACGATGTTATGAAGGAGAAAGGCGTCAACGTTATACCCTACCTGCGG 1083
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        398 AspAspThrSerTyrAspValMetLysGluLysGlyValAsnValIleProTyrLeuArg 417

Qy       1084 CAATCGTGGGTTGATTTGGCGGATAAGTATATGGTAGAGGCACGGTGGTTCTACGGCGGG 1143
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        418 GlnSerTrpValAspLeuAlaAspLysTyrMetValGluAlaArgTrpPheTyrGlyGly 437

Qy       1144 CACAAACCAAGTTTGGAAGAGTATTTGGAGAACTCATGGCAGTCGATAAGTGGGCCCTGT 1203
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        438 HisLysProSerLeuGluGluTyrLeuGluAsnSerTrpGlnSerIleSerGlyProCys 457

Qy       1204 ATGTTAACGCACATATTCTTCCGAGTAACAGATTCGTTCACAAAGGAGACCGTCGACAGT 1263
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        458 MetLeuThrHisIlePhePheArgValThrAspSerPheThrLysGluThrValAspSer 477

Qy       1264 TTGTACAAATACCACGATTTAGTTCGTTGGTCATCCTTCGTTCTGCGGCTTGCTGATGAT 1323
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        478 LeuTyrLysTyrHisAspLeuValArgTrpSerSerPheValLeuArgLeuAlaAspAsp 497

Qy       1324 TTGGGAACCTCGGTGGAAGAGGTGAGCAGAGGGGATGTGCCGAAATCACTTCAGTGCTAC 1383
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        498 LeuGlyThrSerValGluGluValSerArgGlyAspValProLysSerLeuGlnCysTyr 517

Qy       1384 ATGAGTGACTACAATGCATCGGAGGCGGAGGCGCGGAAGCACGTGAAATGGCTGATAGCG 1443
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        518 MetSerAspTyrAsnAlaSerGluAlaGluAlaArgLysHisValLysTrpLeuIleAla 537

Qy       1444 GAGGTGTGGAAGAAGATGAATGCGGAGAGGGTGTCGAAGGATTCTCCATTCGGCAAAGAT 1503
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db         538 GluValTrpLysLysMetAsnAlaGluArgValSerLysAspSerProPheGlyLysAsp 557

Qy        1504 TTTATAGGATGTGCAGTTGATTTAGGAAGGATGGCGCAGTTGATGTACCATAATGGAGAT 1563
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         558 PheIleGlyCysAlaValAspLeuGlyArgMetAlaGlnLeuMetTyrHisAsnGlyAsp 577

Qy        1564 GGGCACGGCACACAACACCCTATTATACATCAACAAATGACCAGAACCTTATTCGAGCCC 1623
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         578 GlyHisGlyThrGlnHisProIleIleHisGlnGlnMetThrArgThrLeuPheGluPro 597

Qy        1624 TTTGCA 1629
               ||||||
Db         598 PheAla 599

RESULT 3
AAY90838
ID   AAY90838 standard; protein; 599 AA.
XX
AC   AAY90838;
XX
DT   25-AUG-2000  (first entry)
XX
DE   Spearmint limonene synthase protein sequence SEQ ID NO:22.
XX
KW   Synthase; protein co-ordinate data; active site; modification; terpenoid;
KW   3-dimensional coordinate; alpha carbon atom; plant; terpene synthase;
KW   isoprenoid; breeding programme; fragrance; flavour; pheromone;
KW   defensive agent; pigment; antitumour; steroid hormone;
KW   signal transduction pathway; bile acid; affinity purification;
KW   photoreceptor; enzymatic synthesis; nutrient supplement;
KW   immunological reagent.
XX
OS   Mentha spicata.
XX
PN   WO200017327-A2.
XX
PD   30-MAR-2000.
XX
PF   17-SEP-1999;   99WO-US021419.
XX
PR   18-SEP-1998;   98US-0100993P.
PR   22-APR-1999;   99US-0130628P.
PR   23-AUG-1999;   99US-0150262P.
XX
PA   (KENT ) UNIV KENTUCKY RES DEPT.
PA   (SALK ) SALK INST BIOLOGICAL STUDIES.
XX
PI   Chappell J,  Manna KR,  Noel JP,  Starks CM;
XX
DR   WPI; 2000-292839/25.
DR   N-PSDB; AAA38923.
XX
PT   Novel terpene synthase enzymes, useful for producing terpene
PT   hydrocarbons, e.g. fragrances or antitumor agents, are derived from known
PT   enzymes by specific amino acid alterations.
XX
PS   Claim 103; Page 371-373; 450pp; English.
XX
CC   The present invention describes an isolated terpene synthase (I)
CC   comprising a region with at least 20% identity to region 265-535 of a 548
CC   amino acid (aa) sequence (Ia), given in AAY90831. (I) contains nine alpha
CC   -carbon atoms (alphaC) that have interatomic distances, between each
CC   other, within tabulated ranges, have a centre point (within a sphere of
CC   radius 2.3 Angstrom ) within tabulated ranges, and have an ordered
CC   arrangement of R groups (defining aa side chains), excluding specific
CC   tabulated arrangements (tables given in the specification). (I), and
CC   related enzymes, are used to produce a wide range of terpenoids (e.g.
CC   cyclic, acylic, optionally hydroxylated), useful e.g. as fragrances,
CC   flavours, pheromones, defensive agents, pigments, antitumour agents,
CC   components of signal transduction pathways, precursors of steroid
CC   hormones and bile acids, as photoreceptors and as co-factor side chains.
CC   Some synthases with little or no catalytic activity (and nucleic acids
CC   encoding them) are used as controls in the analysis of products formed by
CC   enzymatic synthesis; as nutrient supplements; for affinity purification
CC   of isoprenoids; or to develop immunological reagents or nucleic acids for
CC   monitoring expression of terpene synthase or inheritance of the gene in
CC   plant breeding programs. The new synthases may produce novel terpene
CC   products. AAA38910 to AAA38938 and AAY90831 to AAY90859 represent
CC   sequences used in the exemplification of the present invention
XX
SQ   Sequence 599 AA;

Alignment Scores:
Pred. No.:              3.56e-266      Length:         599
Score:                  2876.00        Matches:        542
Percent Similarity:     100.00%        Conservative:   0
Best Local Similarity:  100.00%        Mismatches:     0
Query Match:            98.90%         Indels:         0
DB:                     3              Gaps:           0

US-09-938-956-6 (1-1632) x AAY90838 (1-599)

Qy           4 AGACGATCCGGAAACTACAACCCTTCTCGTTGGGATGTCAACTTCATCCAATCGCTTCTC 63
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          58 ArgArgSerGlyAsnTyrAsnProSerArgTrpAspValAsnPheIleGlnSerLeuLeu 77

Qy          64 AGTGACTATAAGGAGGACAAACACGTGATTAGGGCTTCTGAGCTGGTCACTTTGGTGAAG 123
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          78 SerAspTyrLysGluAspLysHisValIleArgAlaSerGluLeuValThrLeuValLys 97

Qy         124 ATGGAACTGGAGAAAGAAACGGATCAAATTCGACAACTTGAGTTGATCGATGACTTGCAG 183
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          98 MetGluLeuGluLysGluThrAspGlnIleArgGlnLeuGluLeuIleAspAspLeuGln 117

Qy         184 AGGATGGGGCTGTCCGATCATTTCCAAAATGAGTTCAAAGAAATCTTGTCCTCTATATAT 243
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         118 ArgMetGlyLeuSerAspHisPheGlnAsnGluPheLysGluIleLeuSerSerIleTyr 137

Qy         244 CTCGACCATCACTATTACAAGAACCCTTTTCCAAAAGAAGAAAGGGATCTCTACTCCACA 303
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         138 LeuAspHisHisTyrTyrLysAsnProPheProLysGluGluArgAspLeuTyrSerThr 157

Qy         304 TCTCTTGCATTTAGGCTCCTCAGAGAACATGGTTTTCAAGTCGCACAAGAGGTATTCGAT 363
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         158 SerLeuAlaPheArgLeuLeuArgGluHisGlyPheGlnValAlaGlnGluValPheAsp 177

Qy         364 AGTTTCAAGAACGAGGAGGGTGAGTTCAAAGAAAGCCTTAGCGACGACACCAGAGGATTG 423
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         178 SerPheLysAsnGluGluGlyGluPheLysGluSerLeuSerAspAspThrArgGlyLeu 197

Qy         424 TTGCAACTGTATGAAGCTTCCTTTCTGTTGACGGAAGGCGAAACCACGCTCGAGTCAGCG 483
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         198 LeuGlnLeuTyrGluAlaSerPheLeuLeuThrGluGlyGluThrThrLeuGluSerAla 217

Qy         484 AGGGAATTCGCCACCAAATTTTTGGAGGAAAAAGTGAACGAGGGTGGTGTTGATGGCGAC 543
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         218 ArgGluPheAlaThrLysPheLeuGluGluLysValAsnGluGlyGlyValAspGlyAsp 237

Qy         544 CTTTTAACAAGAATCGCATATTCTTTGGACATCCCTCTTCATTGGAGGATTAAAAGGCCA 603
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         238 LeuLeuThrArgIleAlaTyrSerLeuAspIleProLeuHisTrpArgIleLysArgPro 257

Qy         604 AATGCACCTGTGTGGATCGAATGGTATAGGAAGAGGCCCGACATGAATCCAGTAGTGTTG 663
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         258 AsnAlaProValTrpIleGluTrpTyrArgLysArgProAspMetAsnProValValLeu 277

Qy         664 GAGCTTGCCATACTCGACTTAAATATTGTTCAAGCACAATTTCAAGAAGAGCTCAAAGAA 723
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         278 GluLeuAlaIleLeuAspLeuAsnIleValGlnAlaGlnPheGlnGluGluLeuLysGlu 297

Qy         724 TCCTTCAGGTGGTGGAGAAATACTGGGTTTGTTGAGAAGCTGCCCTTCGCAAGGGATAGA 783
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         298 SerPheArgTrpTrpArgAsnThrGlyPheValGluLysLeuProPheAlaArgAspArg 317

Qy         784 CTGGTGGAATGCTACTTTTGGAATACTGGGATCATCGAGCCACGTCAGCATGCAAGTGCA 843
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        318 LeuValGluCysTyrPheTrpAsnThrGlyIleIleGluProArgGlnHisAlaSerAla 337

Qy        844 AGGATAATGATGGGCAAAGTCAACGCTCTGATTACGGTGATCGATGATATTTATGATGTC 903
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        338 ArgIleMetMetGlyLysValAsnAlaLeuIleThrValIleAspAspIleTyrAspVal 357

Qy        904 TATGGCACCTTAGAAGAACTCGAACAATTCACTGACCTCATTCGAAGATGGGATATAAAC 963
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        358 TyrGlyThrLeuGluGluLeuGluGlnPheThrAspLeuIleArgArgTrpAspIleAsn 377

Qy        964 TCAATCGACCAACTTCCCGATTACATGCAACTGTGCTTTCTTGCACTCAACAACTTCGTC 1023
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        378 SerIleAspGlnLeuProAspTyrMetGlnLeuCysPheLeuAlaLeuAsnAsnPheVal 397

Qy       1024 GATGATACATCGTACGATGTTATGAAGGAGAAAGGCGTCAACGTTATACCCTACCTGCGG 1083
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        398 AspAspThrSerTyrAspValMetLysGluLysGlyValAsnValIleProTyrLeuArg 417

Qy       1084 CAATCGTGGGTTGATTTGGCGGATAAGTATATGGTAGAGGCACGGTGGTTCTACGGCGGG 1143
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        418 GlnSerTrpValAspLeuAlaAspLysTyrMetValGluAlaArgTrpPheTyrGlyGly 437

Qy       1144 CACAAACCAAGTTTGGAAGAGTATTTGGAGAACTCATGGCAGTCGATAAGTGGGCCCTGT 1203
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        438 HisLysProSerLeuGluGluTyrLeuGluAsnSerTrpGlnSerIleSerGlyProCys 457

Qy       1204 ATGTTAACGCACATATTCTTCCGAGTAACAGATTCGTTCACAAAGGAGACCGTCGACAGT 1263
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        458 MetLeuThrHisIlePhePheArgValThrAspSerPheThrLysGluThrValAspSer 477

Qy       1264 TTGTACAAATACCACGATTTAGTTCGTTGGTCATCCTTCGTTCTGCGGCTTGCTGATGAT 1323
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        478 LeuTyrLysTyrHisAspLeuValArgTrpSerSerPheValLeuArgLeuAlaAspAsp 497

Qy       1324 TTGGGAACCTCGGTGGAAGAGGTGAGCAGAGGGGATGTGCCGAAATCACTTCAGTGCTAC 1383
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        498 LeuGlyThrSerValGluGluValSerArgGlyAspValProLysSerLeuGlnCysTyr 517

Qy       1384 ATGAGTGACTACAATGCATCGGAGGCGGAGGCGCGGAAGCACGTGAAATGGCTGATAGCG 1443
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        518 MetSerAspTyrAsnAlaSerGluAlaGluAlaArgLysHisValLysTrpLeuIleAla 537

Qy       1444 GAGGTGTGGAAGAAGATGAATGCGGAGAGGGTGTCGAAGGATTCTCCATTCGGCAAAGAT 1503
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        538 GluValTrpLysLysMetAsnAlaGluArgValSerLysAspSerProPheGlyLysAsp 557

Qy       1504 TTTATAGGATGTGCAGTTGATTTAGGAAGGATGGCGCAGTTGATGTACCATAATGGAGAT 1563
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        558 PheIleGlyCysAlaValAspLeuGlyArgMetAlaGlnLeuMetTyrHisAsnGlyAsp 577

Qy       1564 GGGCACGGCACACAACACCCTATTATACATCAACAAATGACCAGAACCTTATTCGAGCCC 1623
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        578 GlyHisGlyThrGlnHisProIleIleHisGlnGlnMetThrArgThrLeuPheGluPro 597

Qy       1624 TTTGCA 1629
              ||||||
Db        598 PheAla 599
```

RESULT 4
AAY70716

```
ID   AAY70716 standard; protein; 599 AA.
XX
AC   AAY70716;
XX
DT   24-JUL-2000  (first entry)
XX
DE   Mint 4S-limonene synthase.
XX
KW   Mint; 4S-limonene synthase; monoterpene; insecticide; metabolic pathway;
KW   geranyl diphosphate; GPP; carveol; S-linalool; insect resistance;
KW   synthase enzyme; monocot; Diabrotica larvae; gene modification.
XX
OS   Mentha spicata.
XX
PN   WO200022150-A2.
XX
PD   20-APR-2000.
XX
PF   05-OCT-1999;   99WO-US023180.
XX
PR   14-OCT-1998;   98US-00172339.
XX
PA   (PION-) PIONEER HI-BRED INT INC.
XX
PI   Meyer TE,  Yalpani N;
XX
DR   WPI; 2000-317999/27.
DR   N-PSDB; AAZ52305.
XX
PT   Manipulating monoterpene biosynthetic pathways in plants for increasing
PT   monoterpene yield and for creating or enhancing insect resistance
PT   involves transforming the plants with genes encoding the monoterpenes.
XX
PS   Claim 1e; Page 68-69; 82pp; English.
XX
CC   The patent discloses a method of manipulating the metabolic pathway of a
CC   plant cell, by transforming it with nucleotide sequences encoding the
CC   monoterpene synthesis pathway enzymes, like 4S-limonene synthase, geranyl
CC   diphosphate (GPP) synthase, carveol synthase or S-linalool synthase. This
CC   method can be used for creating or enhancing resistance of monocot plants
CC   like corn, to insects such as larvae of Diabrotica sp. and to manipulate
CC   the metabolic pathways involving reactions, downstream of those catalysed
CC   by the monoterpene synthase enzymes. The present sequence is the mint 4S-
CC   limonene synthase enzyme. This sequence has insecticidal activity and is
CC   useful for gene modification
XX
SQ   Sequence 599 AA;
```

```
Alignment Scores:
Pred. No.:              3.56e-266     Length:        599
Score:                  2876.00       Matches:       542
Percent Similarity:     100.00%       Conservative:  0
Best Local Similarity:  100.00%       Mismatches:    0
Query Match:            98.90%        Indels:        0
DB:                     3             Gaps:          0
```

US-09-938-956-6 (1-1632) x AAY70716 (1-599)

```
Qy          4 AGACGATCCGGAAACTACAACCCTTCTCGTTGGGATGTCAACTTCATCCAATCGCTTCTC 63
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         58 ArgArgSerGlyAsnTyrAsnProSerArgTrpAspValAsnPheIleGlnSerLeuLeu 77

Qy         64 AGTGACTATAAGGAGGACAAACACGTGATTAGGGCTTCTGAGCTGGTCACTTTGGTGAAG 123
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         78 SerAspTyrLysGluAspLysHisValIleArgAlaSerGluLeuValThrLeuValLys 97

Qy        124 ATGGAACTGGAGAAAGAAACGGATCAAATTCGACAACTTGAGTTGATCGATGACTTGCAG 183
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         98 MetGluLeuGluLysGluThrAspGlnIleArgGlnLeuGluLeuIleAspAspLeuGln 117

Qy        184 AGGATGGGGCTGTCCGATCATTTCCAAAATGAGTTCAAAGAAATCTTGTCCTCTATATAT 243
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        118 ArgMetGlyLeuSerAspHisPheGlnAsnGluPheLysGluIleLeuSerSerIleTyr 137

Qy        244 CTCGACCATCACTATTACAAGAACCCTTTTCCAAAAGAAGAAAGGGATCTCTACTCCACA 303
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        138 LeuAspHisHisTyrTyrLysAsnProPheProLysGluGluArgAspLeuTyrSerThr 157

Qy        304 TCTCTTGCATTTAGGCTCCTCAGAGAACATGGTTTTCAAGTCGCACAAGAGGTATTCGAT 363
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        158 SerLeuAlaPheArgLeuLeuArgGluHisGlyPheGlnValAlaGlnGluValPheAsp 177

Qy        364 AGTTTCAAGAACGAGGAGGGTGAGTTCAAAGAAAGCCTTAGCGACGACACCAGAGGATTG 423
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        178 SerPheLysAsnGluGluGlyGluPheLysGluSerLeuSerAspAspThrArgGlyLeu 197
```

```
Qy      424 TTGCAACTGTATGAAGCTTCCTTTCTGTTGACGGAAGGCGAAACCACGCTCGAGTCAGCG 483
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      198 LeuGlnLeuTyrGluAlaSerPheLeuLeuThrGluGlyGluThrThrLeuGluSerAla 217

Qy      484 AGGGAATTCGCCACCAAATTTTTGGAGGAAAAAGTGAACGAGGGTGGTGTTGATGGCGAC 543
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      218 ArgGluPheAlaThrLysPheLeuGluGluLysValAsnGluGlyGlyValAspGlyAsp 237

Qy      544 CTTTTAACAAGAATCGCATATTCTTTGGACATCCCTCTTCATTGGAGGATTAAAAGGCCA 603
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      238 LeuLeuThrArgIleAlaTyrSerLeuAspIleProLeuHisTrpArgIleLysArgPro 257

Qy      604 AATGCACCTGTGTGGATCGAATGGTATAGGAAGAGGCCCGACATGAATCCAGTAGTGTTG 663
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      258 AsnAlaProValTrpIleGluTrpTyrArgLysArgProAspMetAsnProValValLeu 277

Qy      664 GAGCTTGCCATACTCGACTTAAATATTGTTCAAGCACAATTTCAAGAAGAGCTCAAAGAA 723
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      278 GluLeuAlaIleLeuAspLeuAsnIleValGlnAlaGlnPheGlnGluGluLeuLysGlu 297

Qy      724 TCCTTCAGGTGGTGGAGAAATACTGGGTTTGTTGAGAAGCTGCCCTTCGCAAGGGATAGA 783
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      298 SerPheArgTrpTrpArgAsnThrGlyPheValGluLysLeuProPheAlaArgAspArg 317

Qy      784 CTGGTGGAATGCTACTTTTGGAATACTGGGATCATCGAGCCACGTCAGCATGCAAGTGCA 843
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      318 LeuValGluCysTyrPheTrpAsnThrGlyIleIleGluProArgGlnHisAlaSerAla 337

Qy      844 AGGATAATGATGGGCAAAGTCAACGCTCTGATTACGGTGATCGATGATATTTATGATGTC 903
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      338 ArgIleMetMetGlyLysValAsnAlaLeuIleThrValIleAspAspIleTyrAspVal 357

Qy      904 TATGGCACCTTAGAAGAACTCGAACAATTCACTGACCTCATTCGAAGATGGGATATAAAC 963
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      358 TyrGlyThrLeuGluGluLeuGluGlnPheThrAspLeuIleArgArgTrpAspIleAsn 377

Qy      964 TCAATCGACCAACTTCCCGATTACATGCAACTGTGCTTTCTTGCACTCAACAACTTCGTC 1023
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      378 SerIleAspGlnLeuProAspTyrMetGlnLeuCysPheLeuAlaLeuAsnAsnPheVal 397

Qy     1024 GATGATACATCGTACGATGTTATGAAGGAGAAAGGCGTCAACGTTATACCCTACCTGCGG 1083
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      398 AspAspThrSerTyrAspValMetLysGluLysGlyValAsnValIleProTyrLeuArg 417

Qy     1084 CAATCGTGGGTTGATTTGGCGGATAAGTATATGGTAGAGGCACGGTGGTTCTACGGCGGG 1143
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      418 GlnSerTrpValAspLeuAlaAspLysTyrMetValGluAlaArgTrpPheTyrGlyGly 437

Qy     1144 CACAAACCAAGTTTGGAAGAGTATTTGGAGAACTCATGGCAGTCGATAAGTGGGCCCTGT 1203
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      438 HisLysProSerLeuGluGluTyrLeuGluAsnSerTrpGlnSerIleSerGlyProCys 457

Qy     1204 ATGTTAACGCACATATTCTTCCGAGTAACAGATTCGTTCACAAAGGAGACCGTCGACAGT 1263
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      458 MetLeuThrHisIlePhePheArgValThrAspSerPheThrLysGluThrValAspSer 477

Qy     1264 TTGTACAAATACCACGATTTAGTTCGTTGGTCATCCTTCGTTCTGCGGCTTGCTGATGAT 1323
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      478 LeuTyrLysTyrHisAspLeuValArgTrpSerSerPheValLeuArgLeuAlaAspAsp 497

Qy     1324 TTGGGAACCTCGGTGGAAGAGGTGAGCAGAGGGGATGTGCCGAAATCACTTCAGTGCTAC 1383
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      498 LeuGlyThrSerValGluGluValSerArgGlyAspValProLysSerLeuGlnCysTyr 517

Qy     1384 ATGAGTGACTACAATGCATCGGAGGCGGAGGCGCGGAAGCACGTGAAATGGCTGATAGCG 1443
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      518 MetSerAspTyrAsnAlaSerGluAlaGluAlaArgLysHisValLysTrpLeuIleAla 537

Qy     1444 GAGGTGTGGAAGAAGATGAATGCGGAGAGGGTGTCGAAGGATTCTCCATTCGGCAAAGAT 1503
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      538 GluValTrpLysLysMetAsnAlaGluArgValSerLysAspSerProPheGlyLysAsp 557

Qy     1504 TTTATAGGATGTGCAGTTGATTTAGGAAGGATGGCGCAGTTGATGTACCATAATGGAGAT 1563
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      558 PheIleGlyCysAlaValAspLeuGlyArgMetAlaGlnLeuMetTyrHisAsnGlyAsp 577

Qy     1564 GGGCACGGCACACAACACCCTATTATACATCAACAAATGACCAGAACCTTATTCGAGCCC 1623
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      578 GlyHisGlyThrGlnHisProIleIleHisGlnGlnMetThrArgThrLeuPheGluPro 597

Qy     1624 TTTGCA 1629
            ||||||
Db      598 PheAla 599

RESULT 5
AAE29781
ID   AAE29781 standard; protein; 599 AA.
XX
AC   AAE29781;
XX
DT   24-FEB-2003  (first entry)
XX
DE   M. spicata limonene synthase.
XX
KW   Insect-assisted cross-pollination; scent biosynthetic enzyme; enzyme;
KW   hybrid seed; limonene synthase.
XX
OS   Mentha spicata.
XX
PN   WO200276189-A1.
XX
PD   03-OCT-2002.
XX
PF   24-FEB-2002; 2002WO-IL000142.
XX
PR   22-MAR-2001; 2001US-0277657P.
XX
PA   (SCEN-) SCENTGENE POLLINATION LTD.
XX
PI   Paldi N;
XX
DR   WPI; 2003-018835/01.
DR   N-PSDB; AAD47174.
XX
PT   Enhancing insect assisted cross-pollination between flowering plants
PT   useful for producing hybrid seeds, comprises co-expressing at least one
PT   scent biosynthetic enzyme in plants and growing them in a vicinity with
PT   pollinating insect.
XX
PS   Disclosure; Page 70-71; 89pp; English.
XX
CC   The invention relates to a method of enhancing insect-assisted cross-
CC   pollination between flowering plants of a single plant species that have
CC   at least 2 different genetic backgrounds (e.g. different cultivars). The
CC   method involves co-expressing in plants at least one scent biosynthetic
CC   enzyme and growing the plants in a cross-pollination vicinity in the
CC   presence of at least one pollinating insect. The method is useful for the
CC   production of hybrid seeds. The present sequence is M. spicata limonene
CC   synthase, a scent biosynthetic enzyme
XX
SQ   Sequence 599 AA;

Alignment Scores:
Pred. No.:              3.56e-266     Length:         599
Score:                  2876.00       Matches:        542
Percent Similarity:     100.00%       Conservative:   0
Best Local Similarity:  100.00%       Mismatches:     0
Query Match:            98.90%        Indels:         0
DB:                     6             Gaps:           0

US-09-938-956-6 (1-1632) x AAE29781 (1-599)

Qy        4 AGACGATCCGGAAACTACAACCCTTCTCGTTGGGATGTCAACTTCATCCAATCGCTTCTC 63
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       58 ArgArgSerGlyAsnTyrAsnProSerArgTrpAspValAsnPheIleGlnSerLeuLeu 77
```

```
Qy         64 AGTGACTATAAGGAGGACAAACACGTGATTAGGGCTTCTGAGCTGGTCACTTTGGTGAAG 123
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         78 SerAspTyrLysGluAspLysHisValIleArgAlaSerGluLeuValThrLeuValLys 97

Qy        124 ATGGAACTGGAGAAAGAAACGGATCAAATTCGACAACTTGAGTTGATCGATGACTTGCAG 183
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         98 MetGluLeuGluLysGluThrAspGlnIleArgGlnLeuGluLeuIleAspAspLeuGln 117

Qy        184 AGGATGGGGCTGTCCGATCATTTCCAAAATGAGTTCAAAGAAATCTTGTCCTCTATATAT 243
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        118 ArgMetGlyLeuSerAspHisPheGlnAsnGluPheLysGluIleLeuSerSerIleTyr 137

Qy        244 CTCGACCATCACTATTACAAGAACCCTTTTCCAAAAGAAGAAAGGGATCTCTACTCCACA 303
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        138 LeuAspHisHisTyrTyrLysAsnProPheProLysGluGluArgAspLeuTyrSerThr 157

Qy        304 TCTCTTGCATTTAGGCTCCTCAGAGAACATGGTTTTCAAGTCGCACAAGAGGTATTCGAT 363
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        158 SerLeuAlaPheArgLeuLeuArgGluHisGlyPheGlnValAlaGlnGluValPheAsp 177

Qy        364 AGTTTCAAGAACGAGGAGGGTGAGTTCAAAGAAAGCCTTAGCGACGACACCAGAGGATTG 423
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        178 SerPheLysAsnGluGluGlyGluPheLysGluSerLeuSerAspAspThrArgGlyLeu 197

Qy        424 TTGCAACTGTATGAAGCTTCCTTTCTGTTGACGGAAGGCGAAACCACGCTCGAGTCAGCG 483
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        198 LeuGlnLeuTyrGluAlaSerPheLeuLeuThrGluGlyGluThrThrLeuGluSerAla 217

Qy        484 AGGGAATTCGCCACCAAATTTTTGGAGGAAAAAGTGAACGAGGGTGGTGTTGATGGCGAC 543
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        218 ArgGluPheAlaThrLysPheLeuGluGluLysValAsnGluGlyGlyValAspGlyAsp 237

Qy        544 CTTTTAACAAGAATCGCATATTCTTTGGACATCCCTCTTCATTGGAGGATTAAAAGGCCA 603
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        238 LeuLeuThrArgIleAlaTyrSerLeuAspIleProLeuHisTrpArgIleLysArgPro 257

Qy        604 AATGCACCTGTGTGGATCGAATGGTATAGGAAGAGGCCCGACATGAATCCAGTAGTGTTG 663
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        258 AsnAlaProValTrpIleGluTrpTyrArgLysArgProAspMetAsnProValValLeu 277

Qy        664 GAGCTTGCCATACTCGACTTAAATATTGTTCAAGCACAATTTCAAGAAGAGCTCAAAGAA 723
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        278 GluLeuAlaIleLeuAspLeuAsnIleValGlnAlaGlnPheGlnGluGluLeuLysGlu 297

Qy        724 TCCTTCAGGTGGTGGAGAAATACTGGGTTTGTTGAGAAGCTGCCCTTCGCAAGGGATAGA 783
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        298 SerPheArgTrpTrpArgAsnThrGlyPheValGluLysLeuProPheAlaArgAspArg 317

Qy        784 CTGGTGGAATGCTACTTTTGGAATACTGGGATCATCGAGCCACGTCAGCATGCAAGTGCA 843
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        318 LeuValGluCysTyrPheTrpAsnThrGlyIleIleGluProArgGlnHisAlaSerAla 337

Qy        844 AGGATAATGATGGGCAAAGTCAACGCTCTGATTACGGTGATCGATGATATTTATGATGTC 903
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        338 ArgIleMetMetGlyLysValAsnAlaLeuIleThrValIleAspAspIleTyrAspVal 357

Qy        904 TATGGCACCTTAGAAGAACTCGAACAATTCACTGACCTCATTCGAAGATGGGATATAAAC 963
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        358 TyrGlyThrLeuGluGluLeuGluGlnPheThrAspLeuIleArgArgTrpAspIleAsn 377

Qy        964 TCAATCGACCAACTTCCCGATTACATGCAACTGTGCTTTCTTGCACTCAACAACTTCGTC 1023
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        378 SerIleAspGlnLeuProAspTyrMetGlnLeuCysPheLeuAlaLeuAsnAsnPheVal 397

Qy       1024 GATGATACATCGTACGATGTTATGAAGGAGAAAGGCGTCAACGTTATACCCTACCTGCGG 1083
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        398 AspAspThrSerTyrAspValMetLysGluLysGlyValAsnValIleProTyrLeuArg 417

Qy       1084 CAATCGTGGGTTGATTTGGCGGATAAGTATATGGTAGAGGCACGGTGGTTCTACGGCGGG 1143
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        418 GlnSerTrpValAspLeuAlaAspLysTyrMetValGluAlaArgTrpPheTyrGlyGly 437

Qy       1144 CACAAACCAAGTTTGGAAGAGTATTTGGAGAACTCATGGCAGTCGATAAGTGGGCCCTGT 1203
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        438 HisLysProSerLeuGluGluTyrLeuGluAsnSerTrpGlnSerIleSerGlyProCys 457

Qy       1204 ATGTTAACGCACATATTCTTCCGAGTAACAGATTCGTTCACAAAGGAGACCGTCGACAGT 1263
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        458 MetLeuThrHisIlePhePheArgValThrAspSerPheThrLysGluThrValAspSer 477

Qy       1264 TTGTACAAATACCACGATTTAGTTCGTTGGTCATCCTTCGTTCTGCGGCTTGCTGATGAT 1323
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        478 LeuTyrLysTyrHisAspLeuValArgTrpSerSerPheValLeuArgLeuAlaAspAsp 497

Qy       1324 TTGGGAACCTCGGTGGAAGAGGTGAGCAGAGGGGATGTGCCGAAATCACTTCAGTGCTAC 1383
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        498 LeuGlyThrSerValGluGluValSerArgGlyAspValProLysSerLeuGlnCysTyr 517

Qy       1384 ATGAGTGACTACAATGCATCGGAGGCGGAGGCGCGGAAGCACGTGAAATGGCTGATAGCG 1443
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        518 MetSerAspTyrAsnAlaSerGluAlaGluAlaArgLysHisValLysTrpLeuIleAla 537

Qy       1444 GAGGTGTGGAAGAAGATGAATGCGGAGAGGGTGTCGAAGGATTCTCCATTCGGCAAAGAT 1503
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        538 GluValTrpLysLysMetAsnAlaGluArgValSerLysAspSerProPheGlyLysAsp 557

Qy       1504 TTTATAGGATGTGCAGTTGATTTAGGAAGGATGGCGCAGTTGATGTACCATAATGGAGAT 1563
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        558 PheIleGlyCysAlaValAspLeuGlyArgMetAlaGlnLeuMetTyrHisAsnGlyAsp 577

Qy       1564 GGGCACGGCACACAACACCCTATTATACATCAACAAATGACCAGAACCTTATTCGAGCCC 1623
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        578 GlyHisGlyThrGlnHisProIleIleHisGlnGlnMetThrArgThrLeuPheGluPro 597

Qy       1624 TTTGCA 1629
              ||||||
Db        598 PheAla 599

RESULT 6
AAB29400
ID   AAB29400 standard; protein; 599 AA.
XX
AC   AAB29400;
XX
DT   09-FEB-2001  (first entry)
XX
DE   Spearmint limonene synthase.
XX
KW   Spearmint limonene synthase; transgenic microorganism;
KW   monoterpene production; geranyl diphosphate synthase activity;
KW   monoterpene synthase.
XX
OS   Mentha spicata.
XX
PN   JP2000245482-A.
XX
PD   12-SEP-2000.
XX
PF   05-MAR-1999;   99JP-00059431.
XX
PR   05-MAR-1999;   99JP-00059431.
XX
PA   (SOZO-) SOZOTEKI SEIBUTSU KOGAKU KENKYUSHO KK.
XX
DR   WPI; 2000-642026/62.
DR   N-PSDB; AAC63810.
XX
PT   New transformed microbe useful for the production of monoterpene.
XX
PS   Example; Page 10-13; 15pp; Japanese.
XX
CC   The invention relates to a transformed microorgansim containing a first
CC   gene encoding a protein with geranyl diphosphate synthase activity, and a
CC   second gene encoding a monoterpene synthase. The microorganism is able to
CC   produce the particular monoterpene whose synthesis is catalysed by the
CC   monoterpene synthase that has been introduced. The invention also relates
```

```
CC   to the production of monoterpenes using the novel microorganism. The
CC   microorganism can be used for the production of various monoterpenes
CC   which are useful in industry. In an exemplification of the invention,
CC   Escherichia coli DH5-alpha was transformed with cDNA encoding spearmint
CC   limonene synthase (AAC63810) and DNA encoding Bacillus stearothermophilus
CC   farnesyl diphosphate synthase (AAC63809). The present sequence represents
CC   spearmint limonene synthase
XX
SQ   Sequence 599 AA;

Alignment Scores:
Pred. No.:            2.75e-261     Length:         599
Score:                2825.00       Matches:        530
Percent Similarity:   99.08%        Conservative:   7
Best Local Similarity: 97.79%       Mismatches:     5
Query Match:          97.15%        Indels:         0
DB:                   3             Gaps:           0

US-09-938-956-6 (1-1632) x AAB29400 (1-599)

Qy        4 AGACGATCCGGAAACTACAACCCTTCTCGTTGGGATGTCAACTTCATCCAATCGCTTCTC 63
              |||||||||||||||||||||||||||||||||||||||     ||||||||||||||
Db       58 ArgArgSerGlyAsnTyrAsnProSerArgTrpAspValGluPheIleGlnSerLeuHis 77

Qy       64 AGTGACTATAAGGAGGACAAACACGTGATTAGGGCTTCTGAGCTGGTCACTTTGGTGAAG 123
            |||||||||:::||||||||||||     ||||||||||||||||||||||||||||||
Db       78 SerAspTyrGluGluAspLysHisAlaIleArgAlaSerGluLeuValThrLeuValLys 97

Qy      124 ATGGAACTGGAGAAAGAAACGGATCAAATTCGACAACTTGAGTTGATCGATGACTTGCAG 183
            ||||||||||||||||||||||||     ||||||||||||||||||||||||||||||
Db       98 MetGluLeuGluLysGluThrAspHisIleArgGlnLeuGluLeuIleAspAspLeuGln 117

Qy      184 AGGATGGGGCTGTCCGATCATTTCCAAAATGAGTTCAAAGAAATCTTGTCCTCTATATAT 243
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      118 ArgMetGlyLeuSerAspHisPheGlnAsnGluPheLysGluIleLeuSerSerIleTyr 137

Qy      244 CTCGACCATCACTATTACAAGAACCCTTTTCCAAAAGAAGAAAGGGATCTCTACTCCACA 303
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      138 LeuAspHisHisTyrTyrLysAsnProPheProLysGluGluArgAspLeuTyrSerThr 157

Qy      304 TCTCTTGCATTTAGGCTCCTCAGAGAACATGGTTTTCAAGTCGCACAAGAGGTATTCGAT 363
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      158 SerLeuAlaPheArgLeuLeuArgGluHisGlyPheGlnValAlaGlnGluValPheAsp 177

Qy      364 AGTTTCAAGAACGAGGAGGGTGAGTTCAAAGAAAGCCTTAGCGACGACACCAGAGGATTG 423
            |||||||||||||||||||||||||||||||||||||||||||||||||:::||||||
Db      178 SerPheLysAsnGluGluGlyGluPheLysGluSerLeuSerAspAspThrLysGlyLeu 197

Qy      424 TTGCAACTGTATGAAGCTTCCTTTCTGTTGACGGAAGGCGAAACCACGCTCGAGTCAGCG 483
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      198 LeuGlnLeuTyrGluAlaSerPheLeuLeuThrGluGlyGluThrThrLeuGluSerAla 217

Qy      484 AGGGAATTCGCCACCAAATTTTTGGAGGAAAAAGTGAACGAGGGTGGTGTTGATGGCGAC 543
            |||||||||||||||||||||||||||||:::||||||||||||||||||||||||||||
Db      218 ArgGluPheAlaThrLysPheLeuGluGluArgValAsnGluGlyGlyValAspGlyAsp 237

Qy      544 CTTTTAACAAGAATCGCATATTCTTTGGACATCCCTCTTCATTGGAGGATTAAAAGGCCA 603
            ||||||||||||||||||||||||||||||||||||||||||||||||:::|||||||
Db      238 LeuLeuThrArgIleAlaTyrSerLeuAspIleProLeuHisTrpArgValLysArgPro 257

Qy      604 AATGCACCTGTGTGGATCGAATGGTATAGGAAGAGGCCCGACATGAATCCAGTAGTGTTG 663
            |||||||||     ||||||||||||||||||||||||||||||||||||||||||||||
Db      258 AsnAlaProAlaTrpIleGluTrpTyrArgLysArgProAspMetAsnProValValLeu 277

Qy      664 GAGCTTGCCATACTCGACTTAAATATTGTTCAAGCACAATTTCAAGAAGAGCTCAAAGAA 723
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      278 GluLeuAlaIleLeuAspLeuAsnIleValGlnAlaGlnPheGlnGluGluLeuLysGlu 297

Qy      724 TCCTTCAGGTGGTGGAGAAATACTGGGTTTGTTGAGAAGCTGCCCTTCGCAAGGGATAGA 783
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      298 SerPheArgTrpTrpArgAsnThrGlyPheValGluLysLeuProPheAlaArgAspArg 317

Qy      784 CTGGTGGAATGCTACTTTTGGAATACTGGGATCATCGAGCCACGTCAGCATGCAAGTGCA 843
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      318 LeuValGluCysTyrPheTrpAsnThrGlyIleIleGluProArgGlnHisAlaSerAla 337

Qy      844 AGGATAATGATGGGCAAAGTCAACGCTCTGATTACGGTGATCGATGATATTTATGATGTC 903
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      338 ArgIleMetMetGlyLysValAsnAlaLeuIleThrValIleAspAspIleTyrAspVal 357

Qy      904 TATGGCACCTTAGAAGAACTCGAACAATTCACTGACCTCATTCGAAGATGGGATATAAAC 963
            ||||||||||||||||||||||||||||:::|||||||||||||||||||||||||:::
Db      358 TyrGlyThrLeuGluGluLeuGluGlnPheThrGluLeuIleArgArgTrpAspIleAsp 377

Qy      964 TCAATCGACCAACTTCCCGATTACATGCAACTGTGCTTTCTTGCACTCAACAACTTCGTC 1023
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      378 SerIleAspGlnLeuProAspTyrMetGlnLeuCysPheLeuAlaLeuAsnAsnPheVal 397

Qy     1024 GATGATACATCGTACGATGTTATGAAGGAGAAAGGCGTCAACGTTATACCCTACCTGCGG 1083
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      398 AspAspThrSerTyrAspValMetLysGluLysGlyValAsnValIleProTyrLeuArg 417

Qy     1084 CAATCGTGGGTTGATTTGGCGGATAAGTATATGGTAGAGGCACGGTGGTTCTACGGCGGG 1143
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      418 GlnSerTrpValAspLeuAlaAspLysTyrMetValGluAlaArgTrpPheTyrGlyGly 437

Qy     1144 CACAAACCAAGTTTGGAAGAGTATTTGGAGAACTCATGGCAGTCGATAAGTGGGCCCTGT 1203
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      438 HisLysProSerLeuGluGluTyrLeuGluAsnSerTrpGlnSerIleSerGlyProCys 457

Qy     1204 ATGTTAACGCACATATTCTTCCGAGTAACAGATTCGTTCACAAAGGAGACCGTCGACAGT 1263
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      458 MetLeuThrHisIlePhePheArgValThrAspSerPheThrLysGluThrValAspSer 477

Qy     1264 TTGTACAAATACCACGATTTAGTTCGTTGGTCATCCTTCGTTCTGCGGCTTGCTGATGAT 1323
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      478 LeuTyrLysTyrHisAspLeuValArgTrpSerSerPheValLeuArgLeuAlaAspAsp 497

Qy     1324 TTGGGAACCTCGGTGGAAGAGGTGAGCAGAGGGGATGTGCCGAAATCACTTCAGTGCTAC 1383
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      498 LeuGlyThrSerValGluGluValSerArgGlyAspValProLysSerLeuGlnCysTyr 517

Qy     1384 ATGAGTGACTACAATGCATCGGAGGCGGAGGCGCGGAAGCACGTGAAATGGCTGATAGCG 1443
            ||||||||||||:::|||||||||||||||||||||||||||||||||||||||||||||
Db      518 MetSerAspTyrAspAlaSerGluAlaGluAlaArgLysHisValLysTrpLeuIleAla 537

Qy     1444 GAGGTGTGGAAGAAGATGAATGCGGAGAGGGTGTCGAAGGATTCTCCATTCGGCAAAGAT 1503
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      538 GluValTrpLysLysMetAsnAlaGluArgValSerLysAspSerProPheGlyLysAsp 557

Qy     1504 TTTATAGGATGTGCAGTTGATTTAGGAAGGATGGCGCAGTTGATGTACCATAATGGAGAT 1563
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      558 PheIleGlyCysAlaValAspLeuGlyArgMetAlaGlnLeuMetTyrHisAsnGlyAsp 577

Qy     1564 GGGCACGGCACACAACACCCTATTATACATCAACAAATGACCAGAACCTTATTCGAGCCC 1623
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      578 GlyHisGlyThrGlnHisProIleIleHisGlnGlnMetThrArgThrLeuPheGluPro 597

Qy     1624 TTTGCA 1629
            ||||||
Db      598 PheAla 599

RESULT 7
AAY02332
ID   AAY02332 standard; protein; 591 AA.
XX
AC   AAY02332;
XX
DT   08-JUL-1999  (first entry)
XX
DE   Computer generated 1,8-cineole synthase variant.
XX
KW   (+)-bornyl diphosphate synthase; sage; 1,8-cineole synthase;
KW   (+)-sabinene synthase; monoterpenoid production; plant flavour;
KW   plant aroma; plant defence.
XX
```

```
OS   Synthetic.
OS   Salvia officinalis.
XX
PN   WO9915624-A1.
XX
PD   01-APR-1999.
XX
PF   25-SEP-1998;   98WO-US020120.
XX
PR   25-SEP-1997;   97US-00937540.
XX
PA   (UNIW ) UNIV WASHINGTON STATE RES FOUND.
XX
PI   Croteau RB,  Wise ML,  Katahira EJ,  Savage TJ;
XX
DR   WPI; 1999-254698/21.
XX
PT   Nucleic acid encoding monoterpene synthesis enzymes from sage, used e.g.
PT   to modify flavor, aroma, and defense capability.
XX
PS   Example 7; Page 238-240; 259pp; English.
XX
CC   AAY02329-38 represent computer generated variants of the enzyme 1,8-
CC   cineole synthase from sage. The specification also describes sage (+)-
CC   sabinene synthase and (+)-bornyl diphosphate synthase. Expression vectors
CC   containing nucleic acids encoding these enzymes are used to alter levels
CC   of the specified enzymes in host cells, to increase production of
CC   monoterpenoids, e.g. to modify flavour or aroma of plants, to improve
CC   their defence capability or to alter other interactions mediated by
CC   monoterpenes or produce the enzymes for subsequent use
XX
SQ   Sequence 591 AA;

Alignment Scores:
Pred. No.:              2.07e-147     Length:         591
Score:                  1636.50       Matches:        306
Percent Similarity:     75.96%        Conservative:   108
Best Local Similarity:  56.15%        Mismatches:     116
Query Match:            56.28%        Indels:         15
DB:                     2             Gaps:           7

US-09-938-956-6 (1-1632) x AAY02332 (1-591)

Qy        4 AGACGATCCGGAAACTACAACCCTTCTCGTTGGGATGTCAACTTCATCCAATCGCTTCTC 63
Db       57 ArgArgThrGlyGlyTyrGlnProThrLeuTrpAspPheSerThrIleGlnLeuPheAsp 76

Qy       64 AGTGACTATAAGGAGGACAAACACGTGATTAGGGCTTCTGAGCTGGTCACTTTGGTGAAG 123
Db       77 SerGluTyrLysGluGluLysHisValMetArgAlaAlaGlyMetIleAlaGlnValAsn 96

Qy      124 ATGGAACTGGAGAAAGAAACGGATCAAATTCGACAACTTGAGTTGATCGATGACTTGCAG 183
Db       97 MetLeuLeuGlnGluGluValAspSerIleGlnArgLeuGluLeuIleAspAspLeuArg 116

Qy      184 AGGATGGGGCTGTCCGATCATTTCCAAAATGAGTTCAAAGAAATCTTGTCCTCTATATAT 243
Db      117 ArgLeuGlyIleSerCysHisPheAspArgGluIleValGluIleLeuAsnSerLysTyr 136

Qy      244 CTCGACCATCACTATTACAAGAACCCTTTTCCAAAAGAAGAAAGGGATCTCTACTCCACA 303
Db      137 ------------TyrThrAsnAsnGluIle------AspGluSerAspLeuTyrSerThr 150

Qy      304 TCTCTTGCATTTAGGCTCCTCAGAGAACATGGTTTTCAAGTCGCACAAGAGGTATTCGAT 363
Db      151 AlaLeuArgPheLysLeuLeuArgGlnTyrAspPheSerValSerGlnGluValPheAsp 170

Qy      364 AGTTTCAAGAACGAGGAGGGT---GAGTTCAAAGAAAGCCTTAGCGACGACACCAGAGGA 420
Db      171 CysPheLysAsnAspLysGlyThrAspPheLysProSerLeuValAspAspThrArgGly 190

Qy      421 TTGTTGCAACTGTATGAAGCTTCCTTTCTGTTGACGGAAGGCGAAACCACGCTCGAGTCA 480
Db      191 LeuLeuGlnLeuTyrGluAlaSerPheLeuSerAlaGlnGlyGluGluThrLeuHisLeu 210

Qy      481 GCGAGGGAATTCGCCACCAAATTTTTGGAGGAAAAAGTGAACGAGGGTGGTGTTGATGGC 540
Db      211 AlaArgAspPheAlaThrLysPheLeuHisLysArgValLeu---------ValAspLys 227

Qy      541 GAC------CTTTTAACAAGAATCGCATATTCTTTGGACATCCCTCTTCATTGGAGGATT 594
Db      228 AspIleAsnLeuLeuSerSerIleGluArgAlaLeuGluLeuProThrHisTrpArgVal 247

Qy      595 AAAAGGCCAAATGCACCTGTGTGGATCGAATGGTATAGGAAGAGGCCCGACATGAATCCA 654
Db      248 GlnMetProAsnAlaArgSerPheIleAspAlaTyrLysArgArgProAspMetAsnPro 267

Qy      655 GTAGTGTTGGAGCTTGCCATACTCGACTTAAATATTGTTCAAGCACAATTTCAAGAAGAG 714
Db      268 ThrValLeuGluLeuAlaLysLeuAspPheAsnMetValGlnAlaGlnPheGlnGlnGlu 287

Qy      715 CTCAAAGAATCCTTCAGGTGGTGGAGAAATACTGGGTTTGTTGAGAAGCTGCCCTTCGCA 774
Db      288 LeuLysGluAlaSerArgTrpTrpAsnSerThrGlyLeuValHisGluLeuProPheVal 307

Qy      775 AGGGATAGACTGGTGGAATGCTACTTTTGGAATACTGGGATCATCGAGCCACGTCAGCAT 834
Db      308 ArgAspArgIleValGluCysTyrTyrTrpThrThrGlyValValGluArgArgGluHis 327

Qy      835 GCAAGTGCAAGGATAATGATGGGCAAAGTCAACGCTCTGATTACGGTGATCGATGATATT 894
Db      328 GlyTyrGluArgIleMetLeuThrLysIleAsnAlaLeuValThrThrIleAspAspVal 347

Qy      895 TATGATGTCTATGGCACCTTAGAAGAACTCGAACAATTCACTGACCTCATTCGAAGATGG 954
Db      348 PheAspIleTyrGlyThrLeuGluGluLeuGlnLeuPheThrThrAlaIleGlnArgTrp 367

Qy      955 GATATAAACTCAATCGACCAACTTCCCGATTACATGCAACTGTGCTTTCTTGCACTCAAC 1014
Db      368 AspIleGluSerMetLysGlnLeuProProTyrMetGlnIleCysTyrLeuAlaLeuPhe 387

Qy     1015 AACTTCGTCGATGATACATCGTACGATGTTATGAAGGAGAAAGGCGTCAACGTTATACCC 1074
Db      388 AsnPheValAsnGluMetAlaTyrAspThrLeuArgAspLysGlyPheAsnSerThrPro 407

Qy     1075 TACCTGCGGCAATCGTGGGTTGATTTGGCGGATAAGTATATGGTAGAGGCACGGTGGTTC 1134
Db      408 TyrLeuArgLysAlaTrpValAspLeuValGluSerTyrLeuIleGluAlaLysTrpTyr 427

Qy     1135 TACGGCGGGCACAAACCAAGTTTGGAAGAGTATTTGGAGAACTCATGGCAGTCGATAAGT 1194
Db      428 TyrMetGlyHisLysProSerLeuGluGluTyrMetLysAsnSerTrpIleSerIleGly 447

Qy     1195 GGGCCCTGTATGTTAACGCACATATTCTTCCGAGTAACAGATTCGTTCACAAAGGAGACC 1254
Db      448 GlyIleProIleLeuSerHisLeuPhePheArgLeuThrAspSerIleGluGluGluAsp 467

Qy     1255 GTCGACAGTTTGTACAAATACCACGATTTAGTTCGTTGGTCATCCTTCGTTCTGCGGCTT 1314
Db      468 AlaGluSerMetHisLysTyrHisAspIleValArgAlaSerCysThrIleLeuArgLeu 487

Qy     1315 GCTGATGATTTGGGAACCTCGGTGGAAGAGGTGAGCAGAGGGGATGTGCCGAAATCACTT 1374
Db      488 AlaAspAspMetGlyThrSerLeuAspGluValGluArgGlyAspValProLysSerVal 507

Qy     1375 CAGTGCTACATGAGTGACTACAATGCATCGGAGGCGGAGGCGCGGAAGCACGTGAAATGG 1434
Db      508 GlnCysTyrMetAsnGluLysAsnAlaSerGluGluGluAlaArgGluHisValArgSer 527

Qy     1435 CTGATAGCGGAGGTGTGGAAGAAGATGAATGCGGAGAGGGTGTCGAAGGATTCTCCATTC 1494
Db      528 LeuIleAspGlnThrTrpLysMetMetAsnLysGluMetMetThr------SerSerPhe 545

Qy     1495 GGCAAAGATTTTATAGGATGTGCAGTTGATTTAGGAAGGATGGCGCAGTTGATGTAC--- 1551
Db      546 SerLysTyrPheValGlnValSerAlaAsnLeuAlaArgMetAlaGlnTrpIleTyrGln 565
```

```
Qy        1552 CATAATGGAGATGGGCACGGCACACAACACCCTATTATACATCAACAAATGACCAGAACC 1611
               |||     |||||   |||   |||||  ::::::::::::    :::
Db         566 HisGluSerAspGlyPheGlyMetGlnHisSerLeuValAsnLysMetLeuArgGlyLeu 585

Qy        1612 TTATTCGAGCCCTTT 1626
               ||||||:::    :::
Db         586 LeuPheAspArgTyr 590

RESULT 8
AAY02335
ID   AAY02335 standard; protein; 591 AA.
XX
AC   AAY02335;
XX
DT   08-JUL-1999  (first entry)
XX
DE   Computer generated 1,8-cineole synthase variant.
XX
KW   (+)-bornyl diphosphate synthase; sage; 1,8-cineole synthase;
KW   (+)-sabinene synthase; monoterpenoid production; plant flavour;
KW   plant aroma; plant defence.
XX
OS   Synthetic.
OS   Salvia officinalis.
XX
PN   WO9915624-A1.
XX
PD   01-APR-1999.
XX
PF   25-SEP-1998;   98WO-US020120.
XX
PR   25-SEP-1997;   97US-00937540.
XX
PA   (UNIW ) UNIV WASHINGTON STATE RES FOUND.
XX
PI   Croteau RB,  Wise ML,  Katahira EJ,  Savage TJ;
XX
DR   WPI; 1999-254698/21.
XX
PT   Nucleic acid encoding monoterpene synthesis enzymes from sage, used e.g.
PT   to modify flavor, aroma, and defense capability.
XX
PS   Example 7; Page 244-246; 259pp; English.
XX
CC   AAY02329-38 represent computer generated variants of the enzyme 1,8-
CC   cineole synthase from sage. The specification also describes sage (+)-
CC   sabinene synthase and (+)-bornyl diphosphate synthase. Expression vectors
CC   containing nucleic acids encoding these enzymes are used to alter levels
CC   of the specified enzymes in host cells, to increase production of
CC   monoterpenoids, e.g. to modify flavour or aroma of plants, to improve
CC   their defence capability or to alter other interactions mediated by
CC   monoterpenes or produce the enzymes for subsequent use
XX
SQ   Sequence 591 AA;

Alignment Scores:
Pred. No.:              3.22e-147      Length:         591
Score:                  1634.50        Matches:        305
Percent Similarity:     75.96%         Conservative:   109
Best Local Similarity:  55.96%         Mismatches:     116
Query Match:            56.21%         Indels:         15
DB:                     2              Gaps:           7

US-09-938-956-6 (1-1632) x AAY02335 (1-591)

Qy           4 AGACGATCCGGAAACTACAACCCTTCTCGTTGGGATGTCAACTTCATCCAATCGCTTCTC 63
               ||||||:::|||   |||   |||:::   ||||||    :::   ||||||
Db          57 ArgArgThrGlyGlyTyrGlnProThrLeuTrpAspPheSerThrIleGlnLeuPheAsp 76

Qy          64 AGTGACTATAAGGAGGACAAACACGTGATTAGGGCTTCTGAGCTGGTCACTTTGGTGAAG 123
               |||:::||||||||:::||||||::::::||||||:::   ::::::        |||
Db          77 SerGluTyrLysGluGluLysHisLeuMetArgAlaAlaGlyMetIleAlaGlnValAsn 96

Qy         124 ATGGAACTGGAGAAAGAAACGGATCAAATTCGACAACTTGAGTTGATCGATGACTTGCAG 183
               |||   |||::::::|||   |||   |||::::::||||||||||||||||||||:::
Db          97 MetValLeuGlnGluGluValAspSerIleGlnArgLeuGluLeuIleAspAspLeuArg 116

Qy         184 AGGATGGGGCTGTCCGATCATTTCCAAAATGAGTTCAAAGAAATCTTGTCCTCTATATAT 243
               |||:::|||:::|||   ||||||     |||      |||||||||:::|||    |||
Db         117 ArgLeuGlyIleSerCysHisPheAspArgGluIleValGluIleLeuAsnSerLysTyr 136

Qy         244 CTCGACCATCACTATTACAAGAACCCTTTTCCAAAAGAAGAAAGGGATCTCTACTCCACA 303
                           |||      |||            :::|||   ||||||||||||
Db         137 ------------TyrThrAsnAsnGluIle------AspGluSerAspLeuTyrSerThr 150

Qy         304 TCTCTTGCATTTAGGCTCCTCAGAGAACATGGTTTTCAAGTCGCACAAGAGGTATTCGAT 363
               :::|||   |||:::||||||||||||::::::   |||   |||:::||||||||||||
Db         151 AlaLeuArgPheLysLeuLeuArgGlnTyrAspPheSerValSerGlnGluValPheAsp 170

Qy         364 AGTTTCAAGAACGAGGAGGGT---GAGTTCAAAGAAAGCCTTAGCGACGACACCAGAGGA 420
                  ||||||||::::::|||   :::||||||   ||||||   ||||||||||||||
Db         171 CysPheLysAsnAspLysGlyThrAspPheLysProSerLeuValAspAspThrArgGly 190

Qy         421 TTGTTGCAACTGTATGAAGCTTCCTTTCTGTTGACGGAAGGCGAAACCACGCTCGAGTCA 480
               |||||||||||||||||||||||||||||||||||      :::||||||   ||||||
Db         191 LeuLeuGlnLeuTyrGluAlaSerPheLeuSerAlaGlnGlyGluThrLeuHisLeu 210

Qy         481 GCGAGGGAATTCGCCACCAAATTTTTGGAGGAAAAAGTGAACGAGGGTGGTGTTGATGGC 540
               ||||||:::||||||||||||||||||      ::::::|||            ||||||
Db         211 AlaArgAspPheAlaThrLysPheLeuHisLysArgValLeu---------ValAspLys 227

Qy         541 GAC------CTTTTAACAAGAATCGCATATTCTTTGGACATCCCTCTTCATTGGAGGATT 594
               |||      ||||||:::   |||      :::|||::::::|||   |||||||||:::
Db         228 AspIleAsnLeuLeuSerSerIleGluArgAlaLeuGluLeuProThrHisTrpArgVal 247

Qy         595 AAAAGGCCAAATGCACCTGTGTGGATCGAATGGTATAGGAAGAGGCCCGACATGAATCCA 654
               :::   |||||||||      :::|||:::   |||::::::||||||||||||||||||
Db         248 GlnMetProAsnAlaArgSerPheIleAspAlaTyrLysArgArgProAspMetAsnPro 267

Qy         655 GTAGTGTTGGAGCTTGCCATACTCGACTTAAATATTGTTCAAGCACAATTTCAAGAAGAG 714
                  |||||||||||||||   ||||||   |||:::||||||||||||||||||:::|||
Db         268 ThrValLeuGluLeuAlaLysLeuAspPheAsnMetValGlnAlaGlnPheGlnGlnGlu 287

Qy         715 CTCAAAGAATCCTTCAGGTGGTGGAGAAATACTGGGTTTGTTGAGAAGCTGCCCTTCGCA 774
               |||||||||:::   |||||||||   :::||||||   |||   :::|||||||||
Db         288 LeuLysGluAlaSerArgTrpTrpAsnSerThrGlyLeuValHisGluLeuProPheVal 307

Qy         775 AGGGATAGACTGGTGGAATGCTACTTTTGGAATACTGGGATCATCGAGCCACGTCAGCAT 834
               |||||||||:::||||||||||||:::|||   ||||||::::::|||   |||:::|||
Db         308 ArgAspArgIleValGluCysTyrTyrTrpThrThrGlyValValGluArgArgGluHis 327

Qy         835 GCAAGTGCAAGGATAATGATGGGCAAAGTCAACGCTCTGATTACGGTGATCGATGATATT 894
                     |||||||||:::   |||:::||||||||||||:::|||   |||||||||:::
Db         328 GlyTyrGluArgIleMetLeuThrLysIleAsnAlaLeuValThrThrIleAspAspVal 347

Qy         895 TATGATGTCTATGGCACCTTAGAAGAACTCGAACAATTCACTGACCTCATTCGAAGATGG 954
               :::|||:::||||||||||||||||||||||:::   ||||||      |||:::||||||
Db         348 PheAspIleTyrGlyThrLeuGluGluLeuGlnLeuPheThrThrAlaIleGlnArgTrp 367

Qy         955 GATATAAACTCAATCGACCAACTTCCCGATTACATGCAACTGTGCTTTCTTGCACTCAAC 1014
               ||||||   |||:::   |||||||||   |||||||||:::|||:::|||||||||
Db         368 AspIleGluSerMetLysGlnLeuProProTyrMetGlnIleCysTyrLeuAlaLeuPhe 387

Qy        1015 AACTTCGTCGATGATACATCGTACGATGTTATGAAGGAGAAAGGCGTCAACGTTATACCC 1074
               |||||||||::::::   :::||||||   :::::::::||||||   |||      |||
Db         388 AsnPheValAsnGluMetAlaTyrAspThrLeuArgAspLysGlyPheAsnSerThrPro 407

Qy        1075 TACCTGCGGCAATCGTGGGTTGATTTGGCGGATAAGTATATGGTAGAGGCACGGTGGTTC 1134
               |||||||||::::::||||||||||||      :::   |||::::::||||||:::|||:::
Db         408 TyrLeuArgLysAlaTrpValAspLeuValGluSerTyrLeuIleGluAlaLysTrpTyr 427

Qy        1135 TACGGCGGGCACAAACCAAGTTTGGAAGAGTATTTGGAGAACTCATGGCAGTCGATAAGT 1194
               |||   ||||||||||||||||||||||||||||::::::|||||||||      ||||||
Db         428 TyrMetGlyHisLysProSerLeuGluGluTyrMetLysAsnSerTrpIleSerIleGly 447
```

```
Qy       1195 GGGCCCTGTATGTTAACGCACATATTCTTCCGAGTAACAGATTCGTTCACAAAGGAGACC 1254
Db        448 GlyIleProIleLeuSerHisLeuPhePheArgLeuThrAspSerIleGluGluGluAsp 467

Qy       1255 GTCGACAGTTTGTACAAATACCACGATTTAGTTCGTTGGTCATCCTTCGTTCTGCGGCTT 1314
Db        468 AlaGluSerMetHisLysTyrHisAspIleValArgAlaSerCysThrIleLeuArgLeu 487

Qy       1315 GCTGATGATTTGGGAACCTCGGTGGAAGAGGTGAGCAGAGGGGATGTGCCGAAATCACTT 1374
Db        488 AlaAspAspMetGlyThrSerLeuAspGluValGluArgGlyAspValProLysSerVal 507

Qy       1375 CAGTGCTACATGAGTGACTACAATGCATCGGAGGCGGAGGCGCGGAAGCACGTGAAATGG 1434
Db        508 GlnCysTyrMetAsnGluLysAsnAlaSerGluGluGluAlaArgGluHisValArgSer 527

Qy       1435 CTGATAGCGGAGGTGTGGAAGAAGATGAATGCGGAGAGGGTGTCGAAGGATTCTCCATTC 1494
Db        528 LeuIleAspGlnThrTrpLysMetMetAsnLysGluMetMetThr------SerSerPhe 545

Qy       1495 GGCAAAGATTTTATAGGATGTGCAGTTGATTTAGGAAGGATGGCGCAGTTGATGTAC--- 1551
Db        546 SerLysTyrPheValGlnValSerAlaAsnLeuAlaArgMetAlaGlnTrpIleTyrGln 565

Qy       1552 CATAATGGAGATGGGCACGGCACACAACACCCTATTATACATCAACAAATGACCAGAACC 1611
Db        566 HisGluSerAspGlyPheGlyMetGlnHisSerLeuValAsnLysMetLeuArgGlyLeu 585

Qy       1612 TTATTCGAGCCCTTT 1626
Db        586 LeuPheAspArgTyr 590
```

RESULT 9
AAY02309
ID   AAY02309 standard; protein; 591 AA.
XX
AC   AAY02309;
XX
DT   08-JUL-1999 (first entry)
XX
DE   A sage 1,8-cineole synthase protein.
XX
KW   (+)-bornyl diphosphate synthase; sage; 1,8-cineole synthase;
KW   (+)-sabinene synthase; monoterpenoid production; plant flavour;
KW   plant aroma; plant defence.
XX
OS   Salvia officinalis.
XX
PN   WO9915624-A1.
XX
PD   01-APR-1999.
XX
PF   25-SEP-1998;   98WO-US020120.
XX
PR   25-SEP-1997;   97US-00937540.
XX
PA   (UNIW) UNIV WASHINGTON STATE RES FOUND.
XX
PI   Croteau RB, Wise ML, Katahira EJ, Savage TJ;
XX
DR   WPI; 1999-254698/21.
DR   N-PSDB; AAX35580, AAX35606, AAX35607, AAX35608, AAX35609, AAX35610,
DR   AAX35611, AAX35612, AAX35613, AAX35614.
XX
PT   Nucleic acid encoding monoterpene synthesis enzymes from sage, used e.g.
PT   to modify flavor, aroma, and defense capability.
XX
PS   Example 7; Page 53-54; 259pp; English.
XX
CC   The present sequence represents the enzyme 1,8-cineole synthase from
CC   sage. The specification also describes sage (+)-bornyl diphosphate
CC   synthase and (+)-sabinene synthase. Expression vectors containing nucleic
CC   acids encoding these enzymes are used to alter levels of the specified
CC   enzymes in host cells, to increase production of monoterpenoids, e.g. to
CC   modify flavour or aroma of plants, to improve their defence capability or
CC   to alter other interactions mediated by monoterpenes or produce the
CC   enzymes for subsequent use
XX
SQ   Sequence 591 AA;

Alignment Scores:			
Pred. No.:	4.02e-147	Length:	591
Score:	1633.50	Matches:	305
Percent Similarity:	75.96%	Conservative:	109
Best Local Similarity:	55.96%	Mismatches:	116
Query Match:	56.17%	Indels:	15
DB:	2	Gaps:	7

US-09-938-956-6 (1-1632) x AAY02309 (1-591)

```
Qy          4 AGACGATCCGGAAACTACAACCCTTCTCGTTGGGATGTCAACTTCATCCAATCGCTTCTC 63
Db         57 ArgArgThrGlyGlyTyrGlnProThrLeuTrpAspPheSerThrIleGlnLeuPheAsp 76

Qy         64 AGTGACTATAAGGAGGACAAACACGTGATTAGGGCTTCTGAGCTGGTCACTTTGGTGAAG 123
Db         77 SerGluTyrLysGluGluLysHisLeuMetArgAlaAlaGlyMetIleAlaGlnValAsn 96

Qy        124 ATGGAACTGGAGAAAGAAACGGATCAAATTCGACAACTTGAGTTGATCGATGACTTGCAG 183
Db         97 MetLeuLeuGlnGluGluValAspSerIleGlnArgLeuGluLeuIleAspAspLeuArg 116

Qy        184 AGGATGGGGCTGTCCGATCATTTCCAAAATGAGTTCAAAGAAATCTTGTCCTCTATATAT 243
Db        117 ArgLeuGlyIleSerCysHisPheAspArgGluIleValGluIleLeuAsnSerLysTyr 136

Qy        244 CTCGACCATCACTATTACAAGAACCCTTTTCCAAAAGAAGAAAGGGATCTCTACTCCACA 303
Db        137 ------------TyrThrAsnAsnGluIle------AspGluSerAspLeuTyrSerThr 150

Qy        304 TCTCTTGCATTTAGGCTCCTCAGAGAACATGGTTTTCAAGTCGCACAAGAGGTATTCGAT 363
Db        151 AlaLeuArgPheLysLeuLeuArgGlnTyrAspPheSerValSerGlnGluValPheAsp 170

Qy        364 AGTTTCAAGAACGAGGAGGGT---GAGTTCAAAGAAAGCCTTAGCGACGACACCAGAGGA 420
Db        171 CysPheLysAsnAspLysGlyThrAspPheLysProSerLeuValAspAspThrArgGly 190

Qy        421 TTGTTGCAACTGTATGAAGCTTCCTTTCTGTTGACGGAAGGCGAAACCACGCTCGAGTCA 480
Db        191 LeuLeuGlnLeuTyrGluAlaSerPheLeuSerAlaGlnGlyGluGluThrLeuHisLeu 210

Qy        481 GCGAGGGAATTCGCCACCAAATTTTTGGAGGAAAAAGTGAACGAGGGTGGTGTTGATGGC 540
Db        211 AlaArgAspPheAlaThrLysPheLeuHisLysArgValLeu---------ValAspLys 227

Qy        541 GAC------CTTTTAACAAGAATCGCATATTCTTTGGACATCCCTCTTCATTGGAGGATT 594
Db        228 AspIleAsnLeuLeuSerSerIleGluArgAlaLeuGluLeuProThrHisTrpArgVal 247

Qy        595 AAAAGGCCAAATGCACCTGTGTGGATCGAATGGTATAGGAAGAGGCCCGACATGAATCCA 654
Db        248 GlnMetProAsnAlaArgSerPheIleAspAlaTyrLysArgArgProAspMetAsnPro 267

Qy        655 GTAGTGTTGGAGCTTGCCATACTCGACTTAAATATTGTTCAAGCACAATTTCAAGAAGAG 714
Db        268 ThrValLeuGluLeuAlaLysLeuAspPheAsnMetValGlnAlaGlnPheGlnGlnGlu 287

Qy        715 CTCAAAGAATCCTTCAGGTGGTGGAGAAATACTGGGTTTGTTGAGAAGCTGCCCTTCGCA 774
Db        288 LeuLysGluAlaSerArgTrpTrpAsnSerThrGlyLeuValHisGluLeuProPheVal 307

Qy        775 AGGGATAGACTGGTGGAATGCTACTTTTGGAATACTGGGATCATCGAGCCACGTCAGCAT 834
Db        308 ArgAspArgIleValGluCysTyrTyrTrpThrThrGlyValValGluArgArgGluHis 327
```

```
Qy        835 GCAAGTGCAAGGATAATGATGGGCAAAGTCAACGCTCTGATTACGGTGATCGATGATATT 894
Db        328 GlyTyrGluArgIleMetLeuThrLysIleAsnAlaLeuValThrThrIleAspAspVal 347

Qy        895 TATGATGTCTATGGCACCTTAGAAGAACTCGAACAATTCACTGACCTCATTCGAAGATGG 954
Db        348 PheAspIleTyrGlyThrLeuGluGluLeuGlnLeuPheThrThrAlaIleGlnArgTrp 367

Qy        955 GATATAAACTCAATCGACCAACTTCCCGATTACATGCAACTGTGCTTTCTTGCACTCAAC 1014
Db        368 AspIleGluSerMetLysGlnLeuProProTyrMetGlnIleCysTyrLeuAlaLeuPhe 387

Qy       1015 AACTTCGTCGATGATACATCGTACGATGTTATGAAGGAGAAAGGCGTCAACGTTATACCC 1074
Db        388 AsnPheValAsnGluMetAlaTyrAspThrLeuArgAspLysGlyPheAsnSerThrPro 407

Qy       1075 TACCTGCGGCAATCGTGGGTTGATTTGGCGGATAAGTATATGGTAGAGGCACGGTGGTTC 1134
Db        408 TyrLeuArgLysAlaTrpValAspLeuValGluSerTyrLeuIleGluAlaLysTrpTyr 427

Qy       1135 TACGGCGGGCACAAACCAAGTTTGGAAGAGTATTTGGAGAACTCATGGCAGTCGATAAGT 1194
Db        428 TyrMetGlyHisLysProSerLeuGluGluTyrMetLysAsnSerTrpIleSerIleGly 447

Qy       1195 GGGCCCTGTATGTTAACGCACATATTCTTCCGAGTAACAGATTCGTTCACAAAGGAGACC 1254
Db        448 GlyIleProIleLeuSerHisLeuPhePheArgLeuThrAspSerIleGluGluGluAsp 467

Qy       1255 GTCGACAGTTTGTACAAATACCACGATTTAGTTCGTTGGTCATCCTTCGTTCTGCGGCTT 1314
Db        468 AlaGluSerMetHisLysTyrHisAspIleValArgAlaSerCysThrIleLeuArgLeu 487

Qy       1315 GCTGATGATTTGGGAACCTCGGTGGAAGAGGTGAGCAGAGGGGATGTGCCGAAATCACTT 1374
Db        488 AlaAspAspMetGlyThrSerLeuAspGluValGluArgGlyAspValProLysSerVal 507

Qy       1375 CAGTGCTACATGAGTGACTACAATGCATCGGAGGCGGAGGCGCGGAAGCACGTGAAATGG 1434
Db        508 GlnCysTyrMetAsnGluLysAsnAlaSerGluGluGluAlaArgGluHisValArgSer 527

Qy       1435 CTGATAGCGGAGGTGTGGAAGAAGATGAATGCGGAGAGGGTGTCGAAGGATTCTCCATTC 1494
Db        528 LeuIleAspGlnThrTrpLysMetMetAsnLysGluMetMetThr------SerSerPhe 545

Qy       1495 GGCAAAGATTTTATAGGATGTGCAGTTGATTTAGGAAGGATGGCGCAGTTGATGTAC--- 1551
Db        546 SerLysTyrPheValGlnValSerAlaAsnLeuAlaArgMetAlaGlnTrpIleTyrGln 565

Qy       1552 CATAATGGAGATGGGCACGGCACACAACACCCTATTATACATCAACAAATGACCAGAACC 1611
Db        566 HisGluSerAspGlyPheGlyMetGlnHisSerLeuValAsnLysMetLeuArgGlyLeu 585

Qy       1612 TTATTCGAGCCCTTT 1626
Db        586 LeuPheAspArgTyr 590
```

```
RESULT 10
AAY02331
ID   AAY02331 standard; protein; 591 AA.
XX
AC   AAY02331;
XX
DT   08-JUL-1999  (first entry)
XX
DE   Computer generated 1,8-cineole synthase variant.
XX
KW   (+)-bornyl diphosphate synthase; sage; 1,8-cineole synthase;
KW   (+)-sabinene synthase; monoterpenoid production; plant flavour;
KW   plant aroma; plant defence.
XX
OS   Synthetic.
OS   Salvia officinalis.
XX
PN   WO9915624-A1.
XX
PD   01-APR-1999.
XX
PF   25-SEP-1998;   98WO-US020120.
XX
PR   25-SEP-1997;   97US-00937540.
XX
PA   (UNIW ) UNIV WASHINGTON STATE RES FOUND.
XX
PI   Croteau RB,  Wise ML,  Katahira EJ,  Savage TJ;
XX
DR   WPI; 1999-254698/21.
XX
PT   Nucleic acid encoding monoterpene synthesis enzymes from sage, used e.g.
PT   to modify flavor, aroma, and defense capability.
XX
PS   Example 7; Page 236-238; 259pp; English.
XX
CC   AAY02329-38 represent computer generated variants of the enzyme 1,8-
CC   cineole synthase from sage. The specification also describes sage (+)-
CC   sabinene synthase and (+)-bornyl diphosphate synthase. Expression vectors
CC   containing nucleic acids encoding these enzymes are used to alter levels
CC   of the specified enzymes in host cells, to increase production of
CC   monoterpenoids, e.g. to modify flavour or aroma of plants, to improve
CC   their defence capability or to alter other interactions mediated by
CC   monoterpenes or produce the enzymes for subsequent use
XX
SQ   Sequence 591 AA;

Alignment Scores:
Pred. No.:               4.02e-147     Length:        591
Score:                   1633.50       Matches:       305
Percent Similarity:      75.96%        Conservative:  109
Best Local Similarity:   55.96%        Mismatches:    116
Query Match:             56.17%        Indels:        15
DB:                      2             Gaps:          7

US-09-938-956-6 (1-1632) x AAY02331 (1-591)
```

```
Qy          4 AGACGATCCGGAAACTACAACCCTTCTCGTTGGGATGTCAACTTCATCCAATCGCTTCTC 63
Db         57 ArgArgThrGlyGlyTyrGlnProThrLeuTrpAspPheSerThrIleGlnValPheAsp 76

Qy         64 AGTGACTATAAGGAGGACAAACACGTGATTAGGGCTTCTGAGCTGGTCACTTTGGTGAAG 123
Db         77 SerGluTyrLysGluGluLysHisLeuMetArgAlaAlaGlyMetIleAlaGlnValAsn 96

Qy        124 ATGGAACTGGAGAAAGAAACGGATCAAATTCGACAACTTGAGTTGATCGATGACTTGCAG 183
Db         97 MetLeuLeuGlnGluGluValAspSerIleGlnArgLeuGluLeuIleAspAspLeuArg 116

Qy        184 AGGATGGGGCTGTCCGATCATTTCCAAAATGAGTTCAAAGAAATCTTGTCCTCTATATAT 243
Db        117 ArgLeuGlyIleSerCysHisPheAspArgGluIleValGluIleLeuAsnSerLysTyr 136

Qy        244 CTCGACCATCACTATTACAAGAACCCTTTTCCAAAAGAAGAAAGGGATCTCTACTCCACA 303
Db        137 ------------TyrThrAsnAsnGluIle------AspGluSerAspLeuTyrSerThr 150

Qy        304 TCTCTTGCATTTAGGCTCCTCAGAGAACATGGTTTTCAAGTCGCACAAGAGGTATTCGAT 363
Db        151 AlaLeuArgPheLysLeuLeuArgGlnTyrAspPheSerValSerGlnGluValPheAsp 170

Qy        364 AGTTTCAAGAACGAGGAGGGT---GAGTTCAAAGAAAGCCTTAGCGACGACACCAGAGGA 420
Db        171 CysPheLysAsnAspLysGlyThrAspPheLysProSerLeuValAspAspThrArgGly 190

Qy        421 TTGTTGCAACTGTATGAAGCTTCCTTTCTGTTGACGGAAGGCGAAACCACGCTCGAGTCA 480
Db        191 LeuLeuGlnLeuTyrGluAlaSerPheLeuSerAlaGlnGlyGluGluThrLeuHisLeu 210

Qy        481 GCGAGGGAATTCGCCACCAAATTTTTGGAGGAAAAAGTGAACGAGGGTGGTGTTGATGGC 540
```

```
              ||||||:::|||||||||||||||||   ::::::|||          |||||
Db      211 AlaArgAspPheAlaThrLysPheLeuHisLysArgValLeu---------ValAspLys 227

Qy      541 GAC------CTTTTAACAAGAATCGCATATTCTTTGGACATCCCTCTTCATTGGAGGATT 594
            |||      ||||||:::  |||      :::|||::::::|||  ||||||||:::
Db      228 AspIleAsnLeuLeuSerSerIleGluArgAlaLeuGluLeuProThrHisTrpArgVal 247

Qy      595 AAAAGGCCAAATGCACCTGTGTGGATCGAATGGTATAGGAAGAGGCCCGACATGAATCCA 654
            :::   ||||||||      :::|||:::   |||::::::|||||||||||||||
Db      248 GlnMetProAsnAlaArgSerPheIleAspAlaTyrLysArgArgProAspMetAsnPro 267

Qy      655 GTAGTGTTGGAGCTTGCCATACTCGACTTAAATATTGTTCAAGCACAATTTCAAGAAGAG 714
               ||||||||||||||   ||||||   |||:::||||||||||||||||:::|||
Db      268 ThrValLeuGluLeuAlaLysLeuAspPheAsnMetValGlnAlaGlnPheGlnGlnGlu 287

Qy      715 CTCAAAGAATCCTTCAGGTGGTGGAGAAATACTGGGTTTGTTGAGAAGCTGCCCTTCGCA 774
            |||||||||:::   ||||||||   :::||||||   |||   :::||||||||
Db      288 LeuLysGluAlaSerArgTrpTrpAsnSerThrGlyLeuValHisGluLeuProPheVal 307

Qy      775 AGGGATAGACTGGTGGAATGCTACTTTTGGAATACTGGGATCATCGAGCCACGTCAGCAT 834
            |||||||||:::||||||||||||:::|||   ||||||::::::|||   |||:::|||
Db      308 ArgAspArgIleValGluCysTyrTyrTrpThrThrGlyValValGluArgArgGluHis 327

Qy      835 GCAAGTGCAAGGATAATGATGGGCAAAGTCAACGCTCTGATTACGGTGATCGATGATATT 894
                     |||||||||:::   |||:::||||||||||:::|||   |||||||||:::
Db      328 GlyTyrGluArgIleMetLeuThrLysIleAsnAlaLeuValThrThrIleAspAspVal 347

Qy      895 TATGATGTCTATGGCACCTTAGAAGAACTCGAACAATTCACTGACCTCATTCGAAGATGG 954
            :::|||:::||||||||||||||||||||:::   ||||||      |||:::||||||
Db      348 PheAspIleTyrGlyThrLeuGluGluLeuGlnLeuPheThrThrAlaIleGlnArgTrp 367

Qy      955 GATATAAACTCAATCGACCAACTTCCCGATTACATGCAACTGTGCTTTCTTGCACTCAAC 1014
            ||||||   |||:::   |||||||||   ||||||||||:::|||:::|||||||||
Db      368 AspIleGluSerMetLysGlnLeuProProTyrMetGlnIleCysTyrLeuAlaLeuPhe 387

Qy     1015 AACTTCGTCGATGATACATCGTACGATGTTATGAAGGAGAAAGGCGTCAACGTTATACCC 1074
            |||||||||::::::   :::||||||   :::::::::||||||   |||      |||
Db      388 AsnPheValAsnGluMetAlaTyrAspThrLeuArgAspLysGlyPheAsnSerThrPro 407

Qy     1075 TACCTGCGGCAATCGTGGGTTGATTTGGCGGATAAGTATATGGTAGAGGCACGGTGGTTC 1134
            |||||||||::::::||||||||||||   :::   |||::::::||||||:::|||:::
Db      408 TyrLeuArgLysAlaTrpValAspLeuValGluSerTyrLeuIleGluAlaLysTrpTyr 427

Qy     1135 TACGGCGGGCACAAACCAAGTTTGGAAGAGTATTTGGAGAACTCATGGCAGTCGATAAGT 1194
            |||   |||||||||||||||||||||||||||:::::::|||||||||   ||||||
Db      428 TyrMetGlyHisLysProSerLeuGluGluTyrMetLysAsnSerTrpIleSerIleGly 447

Qy     1195 GGGCCCTGTATGTTAACGCACATATTCTTCCGAGTAACAGATTCGTTCACAAAGGAGACC 1254
            |||      :::|||:::|||:::||||||:::||||||||||         :::|||
Db      448 GlyIleProIleLeuSerHisLeuPhePheArgLeuThrAspSerIleGluGluGluAsp 467

Qy     1255 GTCGACAGTTTGTACAAATACCACGATTTAGTTCGTTGGTCATCCTTCGTTCTGCGGCTT 1314
               :::|||::::::||||||||||||:::||||||   |||      :::|||||||||
Db      468 AlaGluSerMetHisLysTyrHisAspIleValArgAlaSerCysThrIleLeuArgLeu 487

Qy     1315 GCTGATGATTTGGGAACCTCGGTGGAAGAGGTGAGCAGAGGGGATGTGCCGAAATCACTT 1374
            |||||||||:::|||||||||::::::||||||   |||||||||||||||||||||:::
Db      488 AlaAspAspMetGlyThrSerLeuAspGluValGluArgGlyAspValProLysSerVal 507

Qy     1375 CAGTGCTACATGAGTGACTACAATGCATCGGAGGCGGAGGCGCGGAAGCACGTGAAATGG 1434
            ||||||||||||::::::   ||||||||||||   |||||||||:::||||||:::
Db      508 GlnCysTyrMetAsnGluLysAsnAlaSerGluGluGluAlaArgGluHisValArgSer 527

Qy     1435 CTGATAGCGGAGGTGTGGAAGAAGATGAATGCGGAGAGGGTGTCGAAGGATTCTCCATTC 1494
            ||||||   :::   ||||||   ||||||   |||   ::::::      |||   |||
Db      528 LeuIleAspGlnThrTrpLysMetMetAsnLysGluMetMetThr------SerSerPhe 545

Qy     1495 GGCAAAGATTTTATAGGATGTGCAGTTGATTTAGGAAGGATGGCGCAGTTGATGTAC--- 1551
               |||   |||:::      :::   :::|||   ||||||||||||   :::|||
Db      546 SerLysTyrPheValGlnValSerAlaAsnLeuAlaArgMetAlaGlnTrpIleTyrGln 565

Qy     1552 CATAATGGAGATGGGCACGGCACACAACACCCTATTATACATCAACAAATGACCAGAACC 1611
            |||   ||||||   |||   ||||||   ::::::::::::   :::
Db      566 HisGluSerAspGlyPheGlyMetGlnHisSerLeuValAsnLysMetLeuArgGlyLeu 585

Qy     1612 TTATTCGAGCCCTTT 1626
            ||||||:::   :::
Db      586 LeuPheAspArgTyr 590
```

```
RESULT 11
AAY90839
ID   AAY90839 standard; protein; 591 AA.
XX
AC   AAY90839;
XX
DT   25-AUG-2000  (first entry)
XX
DE   Sage 1,8 cineole synthase protein sequence SEQ ID NO:24.
XX
KW   Synthase; protein co-ordinate data; active site; modification; terpenoid;
KW   3-dimensional coordinate; alpha carbon atom; plant; terpene synthase;
KW   isoprenoid; breeding programme; fragrance; flavour; pheromone;
KW   defensive agent; pigment; antitumour; steroid hormone;
KW   signal transduction pathway; bile acid; affinity purification;
KW   photoreceptor; enzymatic synthesis; nutrient supplement;
KW   immunological reagent.
XX
OS   Salvia officinalis.
XX
PN   WO200017327-A2.
XX
PD   30-MAR-2000.
XX
PF   17-SEP-1999;   99WO-US021419.
XX
PR   18-SEP-1998;   98US-0100993P.
PR   22-APR-1999;   99US-0130628P.
PR   23-AUG-1999;   99US-0150262P.
XX
PA   (KENT ) UNIV KENTUCKY RES DEPT.
PA   (SALK ) SALK INST BIOLOGICAL STUDIES.
XX
PI   Chappell J,  Manna KR,  Noel JP,  Starks CM;
XX
DR   WPI; 2000-292839/25.
DR   N-PSDB; AAA38924.
XX
PT   Novel terpene synthase enzymes, useful for producing terpene
PT   hydrocarbons, e.g. fragrances or antitumor agents, are derived from known
PT   enzymes by specific amino acid alterations.
XX
PS   Claim 114; Page 376-377; 450pp; English.
XX
CC   The present invention describes an isolated terpene synthase (I)
CC   comprising a region with at least 20% identity to region 265-535 of a 548
CC   amino acid (aa) sequence (Ia), given in AAY90831. (I) contains nine alpha
CC   -carbon atoms (alphaC) that have interatomic distances, between each
CC   other, within tabulated ranges, have a centre point (within a sphere of
CC   radius 2.3 Angstrom ) within tabulated ranges, and have an ordered
CC   arrangement of R groups (defining aa side chains), excluding specific
CC   tabulated arrangements (tables given in the specification). (I), and
CC   related enzymes, are used to produce a wide range of terpenoids (e.g.
CC   cyclic, acylic, optionally hydroxylated), useful e.g. as fragrances,
CC   flavours, pheromones, defensive agents, pigments, antitumour agents,
CC   components of signal transduction pathways, precursors of steroid
CC   hormones and bile acids, as photoreceptors and as co-factor side chains.
CC   Some synthases with little or no catalytic activity (and nucleic acids
CC   encoding them) are used as controls in the analysis of products formed by
CC   enzymatic synthesis; as nutrient supplements; for affinity purification
CC   of isoprenoids; or to develop immunological reagents or nucleic acids for
CC   monitoring expression of terpene synthase or inheritance of the gene in
CC   plant breeding programs. The new synthases may produce novel terpene
CC   products. AAA38910 to AAA38938 and AAY90831 to AAY90859 represent
CC   sequences used in the exemplification of the present invention
XX
SQ   Sequence 591 AA;
```

```
Alignment Scores:
Pred. No.:              4.02e-147     Length:        591
Score:                  1633.50       Matches:       305
Percent Similarity:     75.96%        Conservative:  109
Best Local Similarity:  55.96%        Mismatches:    116
Query Match:            56.17%        Indels:        15
DB:                     3             Gaps:          7

US-09-938-956-6 (1-1632) x AAY90839 (1-591)

Qy      4 AGACGATCCGGAAACTACAACCCTTCTCGTTGGGATGTCAACTTCATCCAATCGCTTCTC 63
Db     57 ArgArgThrGlyGlyTyrGlnProThrLeuTrpAspPheSerThrIleGlnLeuPheAsp 76

Qy     64 AGTGACTATAAGGAGGACAAACACGTGATTAGGGCTTCTGAGCTGGTCACTTTGGTGAAG 123
Db     77 SerGluTyrLysGluGluLysHisLeuMetArgAlaAlaGlyMetIleAlaGlnValAsn 96

Qy    124 ATGGAACTGGAGAAAGAAACGGATCAAATTCGACAACTTGAGTTGATCGATGACTTGCAG 183
Db     97 MetLeuLeuGlnGluGluValAspSerIleGlnArgLeuGluLeuIleAspAspLeuArg 116

Qy    184 AGGATGGGGCTGTCCGATCATTTCCAAAATGAGTTCAAAGAAATCTTGTCCTCTATATAT 243
Db    117 ArgLeuGlyIleSerCysHisPheAspArgGluIleValGluIleLeuAsnSerLysTyr 136

Qy    244 CTCGACCATCACTATTACAAGAACCCTTTTCCAAAAGAAGAAAGGGATCTCTACTCCACA 303
Db    137 ------------TyrThrAsnAsnGluIle------AspGluSerAspLeuTyrSerThr 150

Qy    304 TCTCTTGCATTTAGGCTCCTCAGAGAACATGGTTTTCAAGTCGCACAAGAGGTATTCGAT 363
Db    151 AlaLeuArgPheLysLeuLeuArgGlnTyrAspPheSerValSerGlnGluValPheAsp 170

Qy    364 AGTTTCAAGAACGAGGAGGGT---GAGTTCAAAGAAAGCCTTAGCGACGACACCAGAGGA 420
Db    171 CysPheLysAsnAspLysGlyThrAspPheLysProSerLeuValAspAspThrArgGly 190

Qy    421 TTGTTGCAACTGTATGAAGCTTCCTTTCTGTTGACGGAAGGCGAAACCACGCTCGAGTCA 480
Db    191 LeuLeuGlnLeuTyrGluAlaSerPheLeuSerAlaGlnGlyGluGluThrLeuHisLeu 210

Qy    481 GCGAGGGAATTCGCCACCAAATTTTTGGAGGAAAAAGTGAACGAGGGTGGTGTTGATGGC 540
Db    211 AlaArgAspPheAlaThrLysPheLeuHisLysArgValLeu---------ValAspLys 227

Qy    541 GAC------CTTTTAACAAGAATCGCATATTCTTTGGACATCCCTCTTCATTGGAGGATT 594
Db    228 AspIleAsnLeuLeuSerSerIleGluArgAlaLeuGluLeuProThrHisTrpArgVal 247

Qy    595 AAAAGGCCAAATGCACCTGTGTGGATCGAATGGTATAGGAAGAGGCCCGACATGAATCCA 654
Db    248 GlnMetProAsnAlaArgSerPheIleAspAlaTyrLysArgArgProAspMetAsnPro 267

Qy    655 GTAGTGTTGGAGCTTGCCATACTCGACTTAAATATTGTTCAAGCACAATTTCAAGAAGAG 714
Db    268 ThrValLeuGluLeuAlaLysLeuAspPheAsnMetValGlnAlaGlnPheGlnGlnGlu 287

Qy    715 CTCAAAGAATCCTTCAGGTGGTGGAGAAATACTGGGTTTGTTGAGAAGCTGCCCTTCGCA 774
Db    288 LeuLysGluAlaSerArgTrpTrpAsnSerThrGlyLeuValHisGluLeuProPheVal 307

Qy    775 AGGGATAGACTGGTGGAATGCTACTTTTGGAATACTGGGATCATCGAGCCACGTCAGCAT 834
Db    308 ArgAspArgIleValGluCysTyrTyrTrpThrThrGlyValValGluArgArgGluHis 327

Qy    835 GCAAGTGCAAGGATAATGATGGGCAAAGTCAACGCTCTGATTACGGTGATCGATGATATT 894
Db    328 GlyTyrGluArgIleMetLeuThrLysIleAsnAlaLeuValThrThrIleAspAspVal 347

Qy    895 TATGATGTCTATGGCACCTTAGAAGAACTCGAACAATTCACTGACCTCATTCGAAGATGG 954
Db    348 PheAspIleTyrGlyThrLeuGluGluLeuGlnLeuPheThrThrAlaIleGlnArgTrp 367

Qy    955 GATATAAACTCAATCGACCAACTTCCCGATTACATGCAACTGTGCTTTCTTGCACTCAAC 1014
Db    368 AspIleGluSerMetLysGlnLeuProProTyrMetGlnIleCysTyrLeuAlaLeuPhe 387

Qy   1015 AACTTCGTCGATGATACATCGTACGATGTTATGAAGGAGAAAGGCGTCAACGTTATACCC 1074
Db    388 AsnPheValAsnGluMetAlaTyrAspThrLeuArgAspLysGlyPheAsnSerThrPro 407

Qy   1075 TACCTGCGGCAATCGTGGGTTGATTTGGCGGATAAGTATATGGTAGAGGCACGGTGGTTC 1134
Db    408 TyrLeuArgLysAlaTrpValAspLeuValGluSerTyrLeuIleGluAlaLysTrpTyr 427

Qy   1135 TACGGCGGGCACAAACCAAGTTTGGAAGAGTATTTGGAGAACTCATGGCAGTCGATAAGT 1194
Db    428 TyrMetGlyHisLysProSerLeuGluGluTyrMetLysAsnSerTrpIleSerIleGly 447

Qy   1195 GGGCCCTGTATGTTAACGCACATATTCTTCCGAGTAACAGATTCGTTCACAAAGGAGACC 1254
Db    448 GlyIleProIleLeuSerHisLeuPhePheArgLeuThrAspSerIleGluGluGluAsp 467

Qy   1255 GTCGACAGTTTGTACAAATACCACGATTTAGTTCGTTGGTCATCCTTCGTTCTGCGGCTT 1314
Db    468 AlaGluSerMetHisLysTyrHisAspIleValArgAlaSerCysThrIleLeuArgLeu 487

Qy   1315 GCTGATGATTTGGGAACCTCGGTGGAAGAGGTGAGCAGAGGGGATGTGCCGAAATCACTT 1374
Db    488 AlaAspAspMetGlyThrSerLeuAspGluValGluArgGlyAspValProLysSerVal 507

Qy   1375 CAGTGCTACATGAGTGACTACAATGCATCGGAGGCGGAGGCGCGGAAGCACGTGAAATGG 1434
Db    508 GlnCysTyrMetAsnGluLysAsnAlaSerGluGluGluAlaArgGluHisValArgSer 527

Qy   1435 CTGATAGCGGAGGTGTGGAAGAAGATGAATGCGGAGAGGGTGTCGAAGGATTCTCCATTC 1494
Db    528 LeuIleAspGlnThrTrpLysMetMetAsnLysGluMetMetThr------SerSerPhe 545

Qy   1495 GGCAAAGATTTTATAGGATGTGCAGTTGATTTAGGAAGGATGGCGCAGTTGATGTAC--- 1551
Db    546 SerLysTyrPheValGlnValSerAlaAsnLeuAlaArgMetAlaGlnTrpIleTyrGln 565

Qy   1552 CATAATGGAGATGGGCACGGCACACAACACCCTATTATACATCAACAAATGACCAGAACC 1611
Db    566 HisGluSerAspGlyPheGlyMetGlnHisSerLeuValAsnLysMetLeuArgGlyLeu 585

Qy   1612 TTATTCGAGCCCTTT 1626
Db    586 LeuPheAspArgTyr 590

RESULT 12
AAY02336
ID   AAY02336 standard; protein; 591 AA.
XX
AC   AAY02336;
XX
DT   08-JUL-1999  (first entry)
XX
DE   Computer generated 1,8-cineole synthase variant.
XX
KW   (+)-bornyl diphosphate synthase; sage; 1,8-cineole synthase;
KW   (+)-sabinene synthase; monoterpenoid production; plant flavour;
KW   plant aroma; plant defence.
XX
OS   Synthetic.
OS   Salvia officinalis.
XX
PN   WO9915624-A1.
XX
PD   01-APR-1999.
XX
PF   25-SEP-1998;    98WO-US020120.
XX
```

```
PR   25-SEP-1997;   97US-00937540.
XX
PA   (UNIW ) UNIV WASHINGTON STATE RES FOUND.
XX
PI   Croteau RB,  Wise ML,  Katahira EJ,  Savage TJ;
XX
DR   WPI; 1999-254698/21.
XX
PT   Nucleic acid encoding monoterpene synthesis enzymes from sage, used e.g.
PT   to modify flavor, aroma, and defense capability.
XX
PS   Example 7; Page 246-248; 259pp; English.
XX
CC   AAY02329-38 represent computer generated variants of the enzyme 1,8-
CC   cineole synthase from sage. The specification also describes sage (+)-
CC   sabinene synthase and (+)-bornyl diphosphate synthase. Expression vectors
CC   containing nucleic acids encoding these enzymes are used to alter levels
CC   of the specified enzymes in host cells, to increase production of
CC   monoterpenoids, e.g. to modify flavour or aroma of plants, to improve
CC   their defence capability or to alter other interactions mediated by
CC   monoterpenes or produce the enzymes for subsequent use
XX
SQ   Sequence 591 AA;

Alignment Scores:
Pred. No.:              5.01e-147      Length:         591
Score:                  1632.50        Matches:        305
Percent Similarity:     75.96%         Conservative:   109
Best Local Similarity:  55.96%         Mismatches:     116
Query Match:            56.14%         Indels:         15
DB:                     2              Gaps:           7

US-09-938-956-6 (1-1632) x AAY02336 (1-591)

Qy          4 AGACGATCCGGAAACTACAACCCTTCTCGTTGGGATGTCAACTTCATCCAATCGCTTCTC 63
                |||||||:::|||   |||   |||:::   ||||||   :::    ||||||
Db         57 ArgArgThrGlyGlyTyrGlnProThrLeuTrpAspPheSerThrIleGlnLeuPheAsp 76

Qy         64 AGTGACTATAAGGAGGACAAACACGTGATTAGGGCTTCTGAGCTGGTCACTTTGGTGAAG 123
                |||:::||||||||:::|||||::::::|||||:::    ::::::       |||
Db         77 SerGluTyrLysGluGluLysHisLeuMetArgAlaAlaGlyMetIleAlaGlnValAsn 96

Qy        124 ATGGAACTGGAGAAAGAAACGGATCAAATTCGACAACTTGAGTTGATCGATGACTTGCAG 183
                |||    |||::::::|||    |||    |||::::::||||||||||||||||||:::
Db         97 MetLeuLeuGlnGluGluLeuAspSerIleGlnArgLeuGluLeuIleAspAspLeuArg 116

Qy        184 AGGATGGGGCTGTCCGATCATTTCCAAAATGAGTTCAAAGAAATCTTGTCCTCTATATAT 243
                |||:::|||:::|||   ||||||         |||     ||||||||:::|||   |||
Db        117 ArgLeuGlyIleSerCysHisPheAspArgGluIleValGluIleLeuAsnSerLysTyr 136

Qy        244 CTCGACCATCACTATTACAAGAACCCTTTTCCAAAAGAAGAAAGGGATCTCTACTCCACA 303
                            |||      |||            :::|||   ||||||||||||||
Db        137 ------------TyrThrAsnAsnGluIle------AspGluSerAspLeuTyrSerThr 150

Qy        304 TCTCTTGCATTTAGGCTCCTCAGAGAACATGGTTTTCAAGTCGCACAAGAGGTATTCGAT 363
                :::|||   |||:::||||||||||::::::    |||   |||:::||||||||||||||
Db        151 AlaLeuArgPheLysLeuLeuArgGlnTyrAspPheSerValSerGlnGluValPheAsp 170

Qy        364 AGTTTCAAGAACGAGGAGGGT---GAGTTCAAAGAAAGCCTTAGCGACGACACCAGAGGA 420
                      |||||||||::::::|||    :::||||||   ||||||   ||||||||||||||
Db        171 CysPheLysAsnAspLysGlyThrAspPheLysProSerLeuValAspAspThrArgGly 190

Qy        421 TTGTTGCAACTGTATGAAGCTTCCTTTCTGTTGACGGAAGGCGAAACCACGCTCGAGTCA 480
                |||||||||||||||||||||||||||||              :::||||||    ||||||
Db        191 LeuLeuGlnLeuTyrGluAlaSerPheLeuSerAlaGlnGlyGluGluThrLeuHisLeu 210

Qy        481 GCGAGGGAATTCGCCACCAAATTTTTGGAGGAAAAAGTGAACGAGGGTGGTGTTGATGGC 540
                ||||||:::||||||||||||||||||    ::::::|||                ||||||
Db        211 AlaArgAspPheAlaThrLysPheLeuHisLysArgValLeu---------ValAspLys 227

Qy        541 GAC------CTTTTAACAAGAATCGCATATTCTTTGGACATCCCTCTTCATTGGAGGATT 594
                |||         ||||||:::   |||         :::|||::::::|||   |||||||||:::
Db        228 AspIleAsnLeuLeuSerSerIleGluArgAlaLeuGluLeuProThrHisTrpArgVal 247

Qy        595 AAAAGGCCAAATGCACCTGTGTGGATCGAATGGTATAGGAAGAGGCCCGACATGAATCCA 654
                :::   |||||||||         :::|||:::   |||::::::|||||||||||||||||
Db        248 GlnMetProAsnAlaArgSerPheIleAspAlaTyrLysArgArgProAspMetAsnPro 267

Qy        655 GTAGTGTTGGAGCTTGCCATACTCGACTTAAATATTGTTCAAGCACAATTTCAAGAAGAG 714
                   ||||||||||||||     ||||||     |||:::||||||||||||||||:::|||
Db        268 ThrValLeuGluLeuAlaLysLeuAspPheAsnMetValGlnAlaGlnPheGlnGlnGlu 287

Qy        715 CTCAAAGAATCCTTCAGGTGGTGGAGAAATACTGGGTTTGTTGAGAAGCTGCCCTTCGCA 774
                |||||||||:::      |||||||||      :::||||||   |||     :::|||||||||
Db        288 LeuLysGluAlaSerArgTrpTrpAsnSerThrGlyLeuValHisGluLeuProPheVal 307

Qy        775 AGGGATAGACTGGTGGAATGCTACTTTTGGAATACTGGGATCATCGAGCCACGTCAGCAT 834
                |||||||||:::|||||||||||||:::|||      ||||||::::::|||    |||:::|||
Db        308 ArgAspArgIleValGluCysTyrTyrTrpThrThrGlyValValGluArgArgGluHis 327

Qy        835 GCAAGTGCAAGGATAATGATGGGCAAAGTCAACGCTCTGATTACGGTGATCGATGATATT 894
                      |||||||||:::      |||:::|||||||||:::|||      |||||||||:::
Db        328 GlyTyrGluArgIleMetLeuThrLysIleAsnAlaLeuValThrThrIleAspAspVal 347

Qy        895 TATGATGTCTATGGCACCTTAGAAGAACTCGAACAATTCACTGACCTCATTCGAAGATGG 954
                :::|||:::||||||||||||||||||||:::      ||||||         |||:::||||||
Db        348 PheAspIleTyrGlyThrLeuGluGluLeuGlnLeuPheThrThrAlaIleGlnArgTrp 367

Qy        955 GATATAAACTCAATCGACCAACTTCCCGATTACATGCAACTGTGCTTTCTTGCACTCAAC 1014
                ||||||   |||:::   |||||||||   |||||||||:::|||:::|||||||||
Db        368 AspIleGluSerMetLysGlnLeuProProTyrMetGlnIleCysTyrLeuAlaLeuPhe 387

Qy       1015 AACTTCGTCGATGATACATCGTACGATGTTATGAAGGAGAAAGGCGTCAACGTTATACCC 1074
                |||||||||::::::   :::||||||   :::::::::||||||   |||      |||
Db        388 AsnPheValAsnGluMetAlaTyrAspThrLeuArgAspLysGlyPheAsnSerThrPro 407

Qy       1075 TACCTGCGGCAATCGTGGGTTGATTTGGCGGATAAGTATATGGTAGAGGCACGGTGGTTC 1134
                |||||||||::::::|||||||||||      :::      |||::::::||||||:::|||:::
Db        408 TyrLeuArgLysAlaTrpValAspLeuValGluSerTyrLeuIleGluAlaLysTrpTyr 427

Qy       1135 TACGGCGGGCACAAACCAAGTTTGGAAGAGTATTTGGAGAACTCATGGCAGTCGATAAGT 1194
                |||   ||||||||||||||||||||||||::::::|||||||||      ||||||
Db        428 TyrMetGlyHisLysProSerLeuGluGluTyrMetLysAsnSerTrpIleSerIleGly 447

Qy       1195 GGGCCCTGTATGTTAACGCACATATTCTTCCGAGTAACAGATTCGTTCACAAAGGAGACC 1254
                |||      :::|||:::|||:::|||||||||:::|||||||||         :::|||
Db        448 GlyIleProIleLeuSerHisLeuPhePheArgLeuThrAspSerIleGluGluGluAsp 467

Qy       1255 GTCGACAGTTTGTACAAATACCACGATTTAGTTCGTTGGTCATCCTTCGTTCTGCGGCTT 1314
                   :::|||::::::||||||||||||:::||||||   |||         :::|||||||||
Db        468 AlaGluSerMetHisLysTyrHisAspIleValArgAlaSerCysThrIleLeuArgLeu 487

Qy       1315 GCTGATGATTTGGGAACCTCGGTGGAAGAGGTGAGCAGAGGGGATGTGCCGAAATCACTT 1374
                |||||||||:::||||||||||||::::::||||||   |||||||||||||||||||:::
Db        488 AlaAspAspMetGlyThrSerLeuAspGluValGluArgGlyAspValProLysSerVal 507

Qy       1375 CAGTGCTACATGAGTGACTACAATGCATCGGAGGCGGAGGCGCGGAAGCACGTGAAATGG 1434
                ||||||||||||::::::   ||||||||||||   |||||||||:::||||||:::
Db        508 GlnCysTyrMetAsnGluLysAsnAlaSerGluGluGluAlaArgGluHisValArgSer 527

Qy       1435 CTGATAGCGGAGGTGTGGAAGAAGATGAATGCGGAGAGGGTGTCGAAGGATTCTCCATTC 1494
                ||||||   :::   ||||||   ||||||   |||   ::::::         |||   |||
Db        528 LeuIleAspGlnThrTrpLysMetMetAsnLysGluMetMetThr------SerSerPhe 545

Qy       1495 GGCAAAGATTTTATAGGATGTGCAGTTGATTTAGGAAGGATGGCGCAGTTGATGTAC--- 1551
                |||   |||:::         :::   :::|||   ||||||||||||   :::|||
Db        546 SerLysTyrPheValGlnValSerAlaAsnLeuAlaArgMetAlaGlnTrpIleTyrGln 565

Qy       1552 CATAATGGAGATGGGCACGGCACACAACACCCTATTATACATCAACAAATGACCAGAACC 1611
                |||      ||||||   |||   ||||||    :::::::::::::      :::
Db        566 HisGluSerAspGlyPheGlyMetGlnHisSerLeuValAsnLysMetLeuArgGlyLeu 585

Qy       1612 TTATTCGAGCCCTTT 1626
                ||||||:::   :::
Db        586 LeuPheAspArgTyr 590
```

```
RESULT 13
AAY02329
ID   AAY02329 standard; protein; 591 AA.
XX
AC   AAY02329;
XX
DT   08-JUL-1999  (first entry)
XX
DE   Computer generated 1,8-cineole synthase variant.
XX
KW   (+)-bornyl diphosphate synthase; sage; 1,8-cineole synthase;
KW   (+)-sabinene synthase; monoterpenoid production; plant flavour;
KW   plant aroma; plant defence.
XX
OS   Synthetic.
OS   Salvia officinalis.
XX
PN   WO9915624-A1.
XX
PD   01-APR-1999.
XX
PF   25-SEP-1998;   98WO-US020120.
XX
PR   25-SEP-1997;   97US-00937540.
XX
PA   (UNIW ) UNIV WASHINGTON STATE RES FOUND.
XX
PI   Croteau RB,  Wise ML,  Katahira EJ,  Savage TJ;
XX
DR   WPI; 1999-254698/21.
XX
PT   Nucleic acid encoding monoterpene synthesis enzymes from sage, used e.g.
PT   to modify flavor, aroma, and defense capability.
XX
PS   Example 7; Page 231-233; 259pp; English.
XX
CC   AAY02329-38 represent computer generated variants of the enzyme 1,8-
CC   cineole synthase from sage. The specification also describes sage (+)-
CC   sabinene synthase and (+)-bornyl diphosphate synthase. Expression vectors
CC   containing nucleic acids encoding these enzymes are used to alter levels
CC   of the specified enzymes in host cells, to increase production of
CC   monoterpenoids, e.g. to modify flavour or aroma of plants, to improve
CC   their defence capability or to alter other interactions mediated by
CC   monoterpenes or produce the enzymes for subsequent use
XX
SQ   Sequence 591 AA;

Alignment Scores:
Pred. No.:              5.01e-147      Length:        591
Score:                  1632.50        Matches:       305
Percent Similarity:     75.96%         Conservative:  109
Best Local Similarity:  55.96%         Mismatches:    116
Query Match:            56.14%         Indels:        15
DB:                     2              Gaps:          7

US-09-938-956-6 (1-1632) x AAY02329 (1-591)

Qy          4 AGACGATCCGGAAACTACAACCCTTCTCGTTGGGATGTCAACTTCATCCAATCGCTTCTC 63
Db         57 ArgArgThrGlyGlyTyrGlnProThrIleTrpAspPheSerThrIleGlnLeuPheAsp 76

Qy         64 AGTGACTATAAGGAGGACAAACACGTGATTAGGGCTTCTGAGCTGGTCACTTTGGTGAAG 123
Db         77 SerGluTyrLysGluGluLysHisLeuMetArgAlaAlaGlyMetIleAlaGlnValAsn 96

Qy        124 ATGGAACTGGAGAAAGAAACGGATCAAATTCGACAACTTGAGTTGATCGATGACTTGCAG 183
Db         97 MetLeuLeuGlnGluGluValAspSerIleGlnArgLeuGluLeuIleAspAspLeuArg 116

Qy        184 AGGATGGGGCTGTCCGATCATTTCCAAAATGAGTTCAAAGAAATCTTGTCCTCTATATAT 243
Db        117 ArgLeuGlyIleSerCysHisPheAspArgGluIleValGluIleLeuAsnSerLysTyr 136

Qy        244 CTCGACCATCACTATTACAAGAACCCTTTTCCAAAAGAAGAAAGGGATCTCTACTCCACA 303
Db        137 ------------TyrThrAsnAsnGluIle------AspGluSerAspLeuTyrSerThr 150

Qy        304 TCTCTTGCATTTAGGCTCCTCAGAGAACATGGTTTTCAAGTCGCACAAGAGGTATTCGAT 363
Db        151 AlaLeuArgPheLysLeuLeuArgGlnTyrAspPheSerValSerGlnGluValPheAsp 170

Qy        364 AGTTTCAAGAACGAGGAGGGT---GAGTTCAAAGAAAGCCTTAGCGACGACACCAGAGGA 420
Db        171 CysPheLysAsnAspLysGlyThrAspPheLysProSerLeuValAspAspThrArgGly 190

Qy        421 TTGTTGCAACTGTATGAAGCTTCCTTTCTGTTGACGGAAGGCGAAACCACGCTCGAGTCA 480
Db        191 LeuLeuGlnLeuTyrGluAlaSerPheLeuSerAlaGlnGlyGluGluThrLeuHisLeu 210

Qy        481 GCGAGGGAATTCGCCACCAAATTTTTGGAGGAAAAAGTGAACGAGGGTGGTGTTGATGGC 540
Db        211 AlaArgAspPheAlaThrLysPheLeuHisLysArgValLeu---------ValAspLys 227

Qy        541 GAC------CTTTTAACAAGAATCGCATATTCTTTGGACATCCCTCTTCATTGGAGGATT 594
Db        228 AspIleAsnLeuLeuSerSerIleGluArgAlaLeuGluLeuProThrHisTrpArgVal 247

Qy        595 AAAAGGCCAAATGCACCTGTGTGGATCGAATGGTATAGGAAGAGGCCCGACATGAATCCA 654
Db        248 GlnMetProAsnAlaArgSerPheIleAspAlaTyrLysArgArgProAspMetAsnPro 267

Qy        655 GTAGTGTTGGAGCTTGCCATACTCGACTTAAATATTGTTCAAGCACAATTTCAAGAAGAG 714
Db        268 ThrValLeuGluLeuAlaLysLeuAspPheAsnMetValGlnAlaGlnPheGlnGlnGlu 287

Qy        715 CTCAAAGAATCCTTCAGGTGGTGGAGAAATACTGGGTTTGTTGAGAAGCTGCCCTTCGCA 774
Db        288 LeuLysGluAlaSerArgTrpTrpAsnSerThrGlyLeuValHisGluLeuProPheVal 307

Qy        775 AGGGATAGACTGGTGGAATGCTACTTTTGGAATACTGGGATCATCGAGCCACGTCAGCAT 834
Db        308 ArgAspArgIleValGluCysTyrTyrTrpThrThrGlyValValGluArgArgGluHis 327

Qy        835 GCAAGTGCAAGGATAATGATGGGCAAAGTCAACGCTCTGATTACGGTGATCGATGATATT 894
Db        328 GlyTyrGluArgIleMetLeuThrLysIleAsnAlaLeuValThrThrIleAspAspVal 347

Qy        895 TATGATGTCTATGGCACCTTAGAAGAACTCGAACAATTCACTGACCTCATTCGAAGATGG 954
Db        348 PheAspIleTyrGlyThrLeuGluGluLeuGlnLeuPheThrThrAlaIleGlnArgTrp 367

Qy        955 GATATAAACTCAATCGACCAACTTCCCGATTACATGCAACTGTGCTTTCTTGCACTCAAC 1014
Db        368 AspIleGluSerMetLysGlnLeuProProTyrMetGlnIleCysTyrLeuAlaLeuPhe 387

Qy       1015 AACTTCGTCGATGATACATCGTACGATGTTATGAAGGAGAAAGGCGTCAACGTTATACCC 1074
Db        388 AsnPheValAsnGluMetAlaTyrAspThrLeuArgAspLysGlyPheAsnSerThrPro 407

Qy       1075 TACCTGCGGCAATCGTGGGTTGATTTGGCGGATAAGTATATGGTAGAGGCACGGTGGTTC 1134
Db        408 TyrLeuArgLysAlaTrpValAspLeuValGluSerTyrLeuIleGluAlaLysTrpTyr 427

Qy       1135 TACGGCGGGCACAAACCAAGTTTGGAAGAGTATTTGGAGAACTCATGGCAGTCGATAAGT 1194
Db        428 TyrMetGlyHisLysProSerLeuGluGluTyrMetLysAsnSerTrpIleSerIleGly 447

Qy       1195 GGGCCCTGTATGTTAACGCACATATTCTTCCGAGTAACAGATTCGTTCACAAAGGAGACC 1254
Db        448 GlyIleProIleLeuSerHisLeuPhePheArgLeuThrAspSerIleGluGluGluAsp 467

Qy       1255 GTCGACAGTTTGTACAAATACCACGATTTAGTTCGTTGGTCATCCTTCGTTCTGCGGCTT 1314
Db        468 AlaGluSerMetHisLysTyrHisAspIleValArgAlaSerCysThrIleLeuArgLeu 487

Qy       1315 GCTGATGATTTGGGAACCTCGGTGGAAGAGGTGAGCAGAGGGGATGTGCCGAAATCACTT 1374
```

```
                 |||||||||:::|||||||||::::::||||||    |||||||||||||||||||||:::
Db        488 AlaAspAspMetGlyThrSerLeuAspGluValGluArgGlyAspValProLysSerVal 507

Qy       1375 CAGTGCTACATGAGTGACTACAATGCATCGGAGGCGGAGGCGCGGAAGCACGTGAAATGG 1434
              |||||||||||::::::   ||||||||||||    |||||||||:::|||||||:::
Db        508 GlnCysTyrMetAsnGluLysAsnAlaSerGluGluGluAlaArgGluHisValArgSer 527

Qy       1435 CTGATAGCGGAGGTGTGGAAGAAGATGAATGCGGAGAGGGTGTCGAAGGATTCTCCATTC 1494
              ||||||    :::    ||||||    ||||||    |||    ::::::    |||    |||
Db        528 LeuIleAspGlnThrTrpLysMetMetAsnLysGluMetMetThr------SerSerPhe 545

Qy       1495 GGCAAAGATTTTATAGGATGTGCAGTTGATTTAGGAAGGATGGCGCAGTTGATGTAC--- 1551
                   |||    |||:::        :::    :::|||    ||||||||||||    :::|||
Db        546 SerLysTyrPheValGlnValSerAlaAsnLeuAlaArgMetAlaGlnTrpIleTyrGln 565

Qy       1552 CATAATGGAGATGGGCACGGCACACAACACCCTATTATACATCAACAAATGACCAGAACC 1611
              |||       ||||||    |||    ||||||    ::::::::::::    :::
Db        566 HisGluSerAspGlyPheGlyMetGlnHisSerLeuValAsnLysMetLeuArgGlyLeu 585

Qy       1612 TTATTCGAGCCCTTT 1626
              |||||||:::    :::
Db        586 LeuPheAspArgTyr 590
```

RESULT 14
AAY02330
ID   AAY02330 standard; protein; 591 AA.
XX
AC   AAY02330;
XX
DT   08-JUL-1999 (first entry)
XX
DE   Computer generated 1,8-cineole synthase variant.
XX
KW   (+)-bornyl diphosphate synthase; sage; 1,8-cineole synthase;
KW   (+)-sabinene synthase; monoterpenoid production; plant flavour;
KW   plant aroma; plant defence.
XX
OS   Synthetic.
OS   Salvia officinalis.
XX
PN   WO9915624-A1.
XX
PD   01-APR-1999.
XX
PF   25-SEP-1998; 98WO-US020120.
XX
PR   25-SEP-1997; 97US-00937540.
XX
PA   (UNIW) UNIV WASHINGTON STATE RES FOUND.
XX
PI   Croteau RB, Wise ML, Katahira EJ, Savage TJ;
XX
DR   WPI; 1999-254698/21.
XX
PT   Nucleic acid encoding monoterpene synthesis enzymes from sage, used e.g.
PT   to modify flavor, aroma, and defense capability.
XX
PS   Example 7; Page 234-235; 259pp; English.
XX
CC   AAY02329-38 represent computer generated variants of the enzyme 1,8-
CC   cineole synthase from sage. The specification also describes sage (+)-
CC   sabinene synthase and (+)-bornyl diphosphate synthase. Expression vectors
CC   containing nucleic acids encoding these enzymes are used to alter levels
CC   of the specified enzymes in host cells, to increase production of
CC   monoterpenoids, e.g. to modify flavour or aroma of plants, to improve
CC   their defence capability or to alter other interactions mediated by
CC   monoterpenes or produce the enzymes for subsequent use
XX
SQ   Sequence 591 AA;

```
Alignment Scores:
Pred. No.:              6.25e-147     Length:        591
Score:                  1631.50       Matches:       304
Percent Similarity:     75.96%        Conservative:  110
Best Local Similarity:  55.78%        Mismatches:    116
Query Match:            56.10%        Indels:        15
DB:                     2             Gaps:          7

US-09-938-956-6 (1-1632) x AAY02330 (1-591)

Qy          4 AGACGATCCGGAAACTACAACCCTTCTCGTTGGGATGTCAACTTCATCCAATCGCTTCTC 63
              |||||||:::|||    |||    |||:::    ||||||    :::    :::|||
Db         57 ArgArgThrGlyGlyTyrGlnProThrLeuTrpAspPheSerThrLeuGlnLeuPheAsp 76

Qy         64 AGTGACTATAAGGAGGACAAACACGTGATTAGGGCTTCTGAGCTGGTCACTTTGGTGAAG 123
              |||:::||||||||||:::||||||::::::||||||:::    ::::::        |||
Db         77 SerGluTyrLysGluGluLysHisLeuMetArgAlaAlaGlyMetIleAlaGlnValAsn 96

Qy        124 ATGGAACTGGAGAAAGAAACGGATCAAATTCGACAACTTGAGTTGATCGATGACTTGCAG 183
              |||    |||::::::|||    |||    |||::::::|||||||||||||||||||||:::
Db         97 MetLeuLeuGlnGluGluValAspSerIleGlnArgLeuGluLeuIleAspAspLeuArg 116

Qy        184 AGGATGGGGCTGTCCGATCATTTCCAAAATGAGTTCAAAGAAATCTTGTCCTCTATATAT 243
              |||:::||||||:::|||    ||||||        |||    ||||||||||:::|||    |||
Db        117 ArgLeuGlyIleSerCysHisPheAspArgGluIleValGluIleLeuAsnSerLysTyr 136

Qy        244 CTCGACCATCACTATTACAAGAACCCTTTTCCAAAAGAAGAAAGGGATCTCTACTCCACA 303
                          |||    |||                :::|||    ||||||||||||
Db        137 ------------TyrThrAsnAsnGluIle------AspGluSerAspLeuTyrSerThr 150

Qy        304 TCTCTTGCATTTAGGCTCCTCAGAGAACATGGTTTTCAAGTCGCACAAGAGGTATTCGAT 363
              :::|||    |||:::||||||||||::::::    |||    |||:::|||||||||||||||
Db        151 AlaLeuArgPheLysLeuLeuArgGlnTyrAspPheSerValSerGlnGluValPheAsp 170

Qy        364 AGTTTCAAGAACGAGGAGGGT---GAGTTCAAAGAAAGCCTTAGCGACGACACCAGAGGA 420
                 |||||||||::::::|||    :::||||||    ||||||    ||||||||||||
Db        171 CysPheLysAsnAspLysGlyThrAspPheLysProSerLeuValAspAspThrArgGly 190

Qy        421 TTGTTGCAACTGTATGAAGCTTCCTTTCTGTTGACGGAAGGCGAAACCACGCTCGAGTCA 480
              ||||||||||||||||||||||||||||||    :::||||||    ||||||
Db        191 LeuLeuGlnLeuTyrGluAlaSerPheLeuSerAlaGlnGlyGluGluThrLeuHisLeu 210

Qy        481 GCGAGGGAATTCGCCACCAAATTTTTGGAGGAAAAAGTGAACGAGGGTGGTGTTGATGGC 540
              ||||||:::||||||||||||||||||    ::::::|||                ||||||
Db        211 AlaArgAspPheAlaThrLysPheLeuHisLysArgValLeu---------ValAspLys 227

Qy        541 GAC------CTTTTAACAAGAATCGCATATTCTTTGGACATCCCTCTTCATTGGAGGATT 594
              |||      ||||||:::    |||    :::||||||:::|||    |||||||||:::
Db        228 AspIleAsnLeuLeuSerSerIleGluArgAlaLeuGluLeuProThrHisTrpArgVal 247

Qy        595 AAAAGGCCAAATGCACCTGTGTGGATCGAATGGTATAGGAAGAGGCCCGACATGAATCCA 654
              :::    |||||||||        :::|||:::    |||::::::||||||||||||||||||
Db        248 GlnMetProAsnAlaArgSerPheIleAspAlaTyrLysArgArgProAspMetAsnPro 267

Qy        655 GTAGTGTTGGAGCTTGCCATACTCGACTTAAATATTGTTCAAGCACAATTTCAAGAAGAG 714
                 |||||||||||||||    ||||||    |||:::||||||||||||||||||:::|||
Db        268 ThrValLeuGluLeuAlaLysLeuAspPheAsnMetValGlnAlaGlnPheGlnGlnGlu 287

Qy        715 CTCAAAGAATCCTTCAGGTGGTGGAGAAATACTGGGTTTGTTGAGAAGCTGCCCTTCGCA 774
              ||||||||:::    |||||||||    :::||||||    |||    :::|||||||||
Db        288 LeuLysGluAlaSerArgTrpTrpAsnSerThrGlyLeuValHisGluLeuProPheVal 307

Qy        775 AGGGATAGACTGGTGGAATGCTACTTTTGGAATACTGGGATCATCGAGCCACGTCAGCAT 834
              ||||||||||:::|||||||||||||::::||    ||||||::::::|||    ||||:::|||
Db        308 ArgAspArgIleValGluCysTyrTyrTrpThrThrGlyValValGluArgArgGluHis 327

Qy        835 GCAAGTGCAAGGATAATGATGGGCAAAGTCAACGCTCTGATTACGGTGATCGATGATATT 894
                       |||||||||:::    |||:::||||||||||:::|||    |||||||||:::
Db        328 GlyTyrGluArgIleMetLeuThrLysIleAsnAlaLeuValThrThrIleAspAspVal 347

Qy        895 TATGATGTCTATGGCACCTTAGAAGAACTCGAACAATTCACTGACCTCATTCGAAGATGG 954
              :::|||:::||||||||||||||||||:::    ||||||        |||:::||||||
Db        348 PheAspIleTyrGlyThrLeuGluGluLeuGlnLeuPheThrThrAlaIleGlnArgTrp 367

Qy        955 GATATAAACTCAATCGACCAACTTCCCGATTACATGCAACTGTGCTTTCTTGCACTCAAC 1014
              ||||||    |||:::    |||||||||    ||||||||||:::|||:::|||||||||
```

```
Db        368 AspIleGluSerMetLysGlnLeuProProTyrMetGlnIleCysTyrLeuAlaLeuPhe 387

Qy       1015 AACTTCGTCGATGATACATCGTACGATGTTATGAAGGAGAAAGGCGTCAACGTTATACCC 1074
              ||||||||::::::    :::||||||   :::::::::||||||   |||       |||
Db        388 AsnPheValAsnGluMetAlaTyrAspThrLeuArgAspLysGlyPheAsnSerThrPro 407

Qy       1075 TACCTGCGGCAATCGTGGGTTGATTTGGCGGATAAGTATATGGTAGAGGCACGGTGGTTC 1134
              |||||||||::::::|||||||||||     :::   |||::::::||||||:::|||:::
Db        408 TyrLeuArgLysAlaTrpValAspLeuValGluSerTyrLeuIleGluAlaLysTrpTyr 427

Qy       1135 TACGGCGGGCACAAACCAAGTTTGGAAGAGTATTTGGAGAACTCATGGCAGTCGATAAGT 1194
              |||   |||||||||||||||||||||||||||||||::::::|||||||||   ||||||
Db        428 TyrMetGlyHisLysProSerLeuGluGluTyrMetLysAsnSerTrpIleSerIleGly 447

Qy       1195 GGGCCCTGTATGTTAACGCACATATTCTTCCGAGTAACAGATTCGTTCACAAAGGAGACC 1254
              |||        :::||||:::||||:::|||||||||:::|||||||||        :::|||
Db        448 GlyIleProIleLeuSerHisLeuPhePheArgLeuThrAspSerIleGluGluGluAsp 467

Qy       1255 GTCGACAGTTTGTACAAATACCACGATTTAGTTCGTTGGTCATCCTTCGTTCTGCGGCTT 1314
                :::|||::::::||||||||||||:::||||||   |||      :::|||||||||
Db        468 AlaGluSerMetHisLysTyrHisAspIleValArgAlaSerCysThrIleLeuArgLeu 487

Qy       1315 GCTGATGATTTGGGAACCTCGGTGGAAGAGGTGAGCAGAGGGGATGTGCCGAAATCACTT 1374
              |||||||||:::||||||||||::::::||||||   |||||||||||||||||||||:::
Db        488 AlaAspAspMetGlyThrSerLeuAspGluValGluArgGlyAspValProLysSerVal 507

Qy       1375 CAGTGCTACATGAGTGACTACAATGCATCGGAGGCGGAGGCGCGGAAGCACGTGAAATGG 1434
              ||||||||||||::::::    ||||||||||||   |||||||||:::||||||:::
Db        508 GlnCysTyrMetAsnGluLysAsnAlaSerGluGluGluAlaArgGluHisValArgSer 527

Qy       1435 CTGATAGCGGAGGTGTGGAAGAAGATGAATGCGGAGAGGGTGTCGAAGGATTCTCCATTC 1494
              ||||||   :::   ||||||   ||||||   |||   ::::::        |||   |||
Db        528 LeuIleAspGlnThrTrpLysMetMetAsnLysGluMetMetThr------SerSerPhe 545

Qy       1495 GGCAAAGATTTTATAGGATGTGCAGTTGATTTAGGAAGGATGGCGCAGTTGATGTAC--- 1551
                 |||   |||:::        :::   :::|||   ||||||||||||   :::|||
Db        546 SerLysTyrPheValGlnValSerAlaAsnLeuAlaArgMetAlaGlnTrpIleTyrGln 565

Qy       1552 CATAATGGAGATGGGCACGGCACACAACACCCTATTATACATCAACAAATGACCAGAACC 1611
              |||      ||||||   |||   ||||||   ::::::::::::   :::
Db        566 HisGluSerAspGlyPheGlyMetGlnHisSerLeuValAsnLysMetLeuArgGlyLeu 585

Qy       1612 TTATTCGAGCCCTTT 1626
              ||||||:::   :::
Db        586 LeuPheAspArgTyr 590
```

RESULT 15
AAY02337
ID   AAY02337 standard; protein; 591 AA.
XX
AC   AAY02337;
XX
DT   08-JUL-1999 (first entry)
XX
DE   Computer generated 1,8-cineole synthase variant.
XX
KW   (+)-bornyl diphosphate synthase; sage; 1,8-cineole synthase;
KW   (+)-sabinene synthase; monoterpenoid production; plant flavour;
KW   plant aroma; plant defence.
XX
OS   Synthetic.
OS   Salvia officinalis.
XX
PN   WO9915624-A1.
XX
PD   01-APR-1999.
XX
PF   25-SEP-1998;   98WO-US020120.
XX
PR   25-SEP-1997;   97US-00937540.
XX
PA   (UNIW) UNIV WASHINGTON STATE RES FOUND.
XX
PI   Croteau RB, Wise ML, Katahira EJ, Savage TJ;
XX
DR   WPI; 1999-254698/21.
XX
PT   Nucleic acid encoding monoterpene synthesis enzymes from sage, used e.g.
PT   to modify flavor, aroma, and defense capability.
XX
PS   Example 7; Page 249-250; 259pp; English.
XX
CC   AAY02329-38 represent computer generated variants of the enzyme 1,8-
CC   cineole synthase from sage. The specification also describes sage (+)-
CC   sabinene synthase and (+)-bornyl diphosphate synthase. Expression vectors
CC   containing nucleic acids encoding these enzymes are used to alter levels
CC   of the specified enzymes in host cells, to increase production of
CC   monoterpenoids, e.g. to modify flavour or aroma of plants, to improve
CC   their defence capability or to alter other interactions mediated by
CC   monoterpenes or produce the enzymes for subsequent use
XX
SQ   Sequence 591 AA;

Alignment Scores:

Pred. No.:	6.25e-147	Length:	591
Score:	1631.50	Matches:	304
Percent Similarity:	75.96%	Conservative:	110
Best Local Similarity:	55.78%	Mismatches:	116
Query Match:	56.10%	Indels:	15
DB:	2	Gaps:	7

US-09-938-956-6 (1-1632) x AAY02337 (1-591)

```
Qy          4 AGACGATCCGGAAACTACAACCCTTCTCGTTGGGATGTCAACTTCATCCAATCGCTTCTC 63
              ||||||:::|||   |||   |||:::   ||||||   :::   ||||||
Db         57 ArgArgThrGlyGlyTyrGlnProThrLeuTrpAspPheSerThrIleGlnLeuPheAsp 76

Qy         64 AGTGACTATAAGGAGGACAAACACGTGATTAGGGCTTCTGAGCTGGTCACTTTGGTGAAG 123
              |||:::||||||||||:::||||||:::::::|||||||:::   ::::::      |||
Db         77 SerGluTyrLysGluGluLysHisLeuMetArgAlaAlaGlyMetIleAlaGlnValAsn 96

Qy        124 ATGGAACTGGAGAAAGAAACGGATCAAATTCGACAACTTGAGTTGATCGATGACTTGCAG 183
              |||   |||::::::|||   |||   :::::::::|||||||||||||||||||||:::
Db         97 MetLeuLeuGlnGluGluValAspSerLeuGlnArgLeuGluLeuIleAspAspLeuArg 116

Qy        184 AGGATGGGGCTGTCCGATCATTTCCAAAATGAGTTCAAAGAAATCTTGTCCTCTATATAT 243
              |||:::|||:::|||   ||||||         |||   ||||||||||:::|||   |||
Db        117 ArgLeuGlyIleSerCysHisPheAspArgGluIleValGluIleLeuAsnSerLysTyr 136

Qy        244 CTCGACCATCACTATTACAAGAACCCTTTTCCAAAAGAAGAAAGGGATCTCTACTCCACA 303
                          |||      |||               :::|||   ||||||||||||||
Db        137 ------------TyrThrAsnAsnGluIle------AspGluSerAspLeuTyrSerThr 150

Qy        304 TCTCTTGCATTTAGGCTCCTCAGAGAACATGGTTTTCAAGTCGCACAAGAGGTATTCGAT 363
              :::|||   |||:::|||||||||::::::   |||   |||:::||||||||||||||||
Db        151 AlaLeuArgPheLysLeuLeuArgGlnTyrAspPheSerValSerGlnGluValPheAsp 170

Qy        364 AGTTTCAAGAACGAGGAGGGT---GAGTTCAAAGAAAGCCTTAGCGACGACACCAGAGGA 420
                 |||||||||::::::|||      :::||||||   ||||||   |||||||||||||||
Db        171 CysPheLysAsnAspLysGlyThrAspPheLysProSerLeuValAspAspThrArgGly 190

Qy        421 TTGTTGCAACTGTATGAAGCTTCCTTTCTGTTGACGGAAGGCGAAACCACGCTCGAGTCA 480
              |||||||||||||||||||||||||||||||         :::||||||   ||||||
Db        191 LeuLeuGlnLeuTyrGluAlaSerPheLeuSerAlaGlnGlyGluGluThrLeuHisLeu 210

Qy        481 GCGAGGGAATTCGCCACCAAATTTTTGGAGGAAAAAGTGAACGAGGGTGGTGTTGATGGC 540
              ||||||:::||||||||||||||||||   ::::::|||                ||||||
Db        211 AlaArgAspPheAlaThrLysPheLeuHisLysArgValLeu---------ValAspLys 227

Qy        541 GAC------CTTTTAACAAGAATCGCATATTCTTTGGACATCCCTCTTCATTGGAGGATT 594
              |||      ||||||:::   |||      :::|||::::::|||   |||||||||:::
Db        228 AspIleAsnLeuLeuSerSerIleGluArgAlaLeuGluLeuProThrHisTrpArgVal 247

Qy        595 AAAAGGCCAAATGCACCTGTGTGGATCGAATGGTATAGGAAGAGGCCCGACATGAATCCA 654
               :::   |||||||||   :::|||:::   |||::::::|||||||||||||||||||
Db        248 GlnMetProAsnAlaArgSerPheIleAspAlaTyrLysArgArgProAspMetAsnPro 267
```

```
Qy        655 GTAGTGTTGGAGCTTGCCATACTCGACTTAAATATTGTTCAAGCACAATTTCAAGAAGAG 714
                 ||||||||||||||     ||||||   |||:::||||||||||||||||:::|||
Db        268 ThrValLeuGluLeuAlaLysLeuAspPheAsnMetValGlnAlaGlnPheGlnGlnGlu 287

Qy        715 CTCAAAGAATCCTTCAGGTGGTGGAGAAATACTGGGTTTGTTGAGAAGCTGCCCTTCGCA 774
              |||||||||:::   |||||||||   :::||||||   |||   :::|||||||||
Db        288 LeuLysGluAlaSerArgTrpTrpAsnSerThrGlyLeuValHisGluLeuProPheVal 307

Qy        775 AGGGATAGACTGGTGGAATGCTACTTTTGGAATACTGGGATCATCGAGCCACGTCAGCAT 834
              ||||||||:::|||||||||||||:::|||   ||||||::::::|||   |||:::|||
Db        308 ArgAspArgIleValGluCysTyrTyrTrpThrThrGlyValValGluArgArgGluHis 327

Qy        835 GCAAGTGCAAGGATAATGATGGGCAAAGTCAACGCTCTGATTACGGTGATCGATGATATT 894
                       |||||||||:::   |||:::|||||||||:::|||   |||||||||:::
Db        328 GlyTyrGluArgIleMetLeuThrLysIleAsnAlaLeuValThrThrIleAspAspVal 347

Qy        895 TATGATGTCTATGGCACCTTAGAAGAACTCGAACAATTCACTGACCTCATTCGAAGATGG 954
              :::|||:::||||||||||||||||||||||:::      ||||||      |||:::||||||
Db        348 PheAspIleTyrGlyThrLeuGluGluLeuGlnLeuPheThrThrAlaIleGlnArgTrp 367

Qy        955 GATATAAACTCAATCGACCAACTTCCCGATTACATGCAACTGTGCTTTCTTGCACTCAAC 1014
              ||||||   |||:::   |||||||||   |||||||||:::|||:::|||||||||
Db        368 AspIleGluSerMetLysGlnLeuProProTyrMetGlnIleCysTyrLeuAlaLeuPhe 387

Qy       1015 AACTTCGTCGATGATACATCGTACGATGTTATGAAGGAGAAAGGCGTCAACGTTATACCC 1074
              |||||||||::::::   :::||||||   :::::::::||||||   |||      |||
Db        388 AsnPheValAsnGluMetAlaTyrAspThrLeuArgAspLysGlyPheAsnSerThrPro 407

Qy       1075 TACCTGCGGCAATCGTGGGTTGATTTGGCGGATAAGTATATGGTAGAGGCACGGTGGTTC 1134
              |||||||||::::::|||||||||||||   :::   |||::::::||||||:::|||:::
Db        408 TyrLeuArgLysAlaTrpValAspLeuValGluSerTyrLeuIleGluAlaLysTrpTyr 427

Qy       1135 TACGGCGGGCACAAACCAAGTTTGGAAGAGTATTTGGAGAACTCATGGCAGTCGATAAGT 1194
              |||   |||||||||||||||||||||||||||::::::|||||||||   ||||||
Db        428 TyrMetGlyHisLysProSerLeuGluGluTyrMetLysAsnSerTrpIleSerIleGly 447

Qy       1195 GGGCCCTGTATGTTAACGCACATATTCTTCCGAGTAACAGATTCGTTCACAAAGGAGACC 1254
              |||      :::|||:::|||:::|||||||||:::|||||||||      :::|||
Db        448 GlyIleProIleLeuSerHisLeuPhePheArgLeuThrAspSerIleGluGluGluAsp 467

Qy       1255 GTCGACAGTTTGTACAAATACCACGATTTAGTTCGTTGGTCATCCTTCGTTCTGCGGCTT 1314
                 :::|||::::::||||||||||||:::||||||   |||      :::|||||||||
Db        468 AlaGluSerMetHisLysTyrHisAspIleValArgAlaSerCysThrIleLeuArgLeu 487

Qy       1315 GCTGATGATTTGGGAACCTCGGTGGAAGAGGTGAGCAGAGGGGATGTGCCGAAATCACTT 1374
              |||||||||:::|||||||||::::::||||||   ||||||||||||||||||:::
Db        488 AlaAspAspMetGlyThrSerLeuAspGluValGluArgGlyAspValProLysSerVal 507

Qy       1375 CAGTGCTACATGAGTGACTACAATGCATCGGAGGCGGAGGCGCGGAAGCACGTGAAATGG 1434
              ||||||||||||::::::   ||||||||||||   |||||||||:::||||||:::
Db        508 GlnCysTyrMetAsnGluLysAsnAlaSerGluGluGluAlaArgGluHisValArgSer 527

Qy       1435 CTGATAGCGGAGGTGTGGAAGAAGATGAATGCGGAGAGGGTGTCGAAGGATTCTCCATTC 1494
              ||||||   :::   ||||||   ||||||   |||   ::::::      |||   |||
Db        528 LeuIleAspGlnThrTrpLysMetMetAsnLysGluMetMetThr------SerSerPhe 545

Qy       1495 GGCAAAGATTTTATAGGATGTGCAGTTGATTTAGGAAGGATGGCGCAGTTGATGTAC--- 1551
                 |||   |||:::      :::   :::|||   ||||||||||||   :::|||
Db        546 SerLysTyrPheValGlnValSerAlaAsnLeuAlaArgMetAlaGlnTrpIleTyrGln 565

Qy       1552 CATAATGGAGATGGGCACGGCACACAACACCCTATTATACATCAACAAATGACCAGAACC 1611
              |||      ||||||   |||   ||||||   ::::::::::::   :::
Db        566 HisGluSerAspGlyPheGlyMetGlnHisSerLeuValAsnLysMetLeuArgGlyLeu 585

Qy       1612 TTATTCGAGCCCTTT 1626
              ||||||:::   :::
Db        586 LeuPheAspArgTyr 590
```

Search completed: February 25, 2004, 05:46:09
Job time : 119.5 secs

```
OM nucleic - protein search, using frame_plus_n2p model

Run on:         February 25, 2004, 05:43:32 ; Search time 24 Seconds
                                           (without alignments)
                                           7021.132 Million cell updates/sec

Title:          US-09-938-956-6
Perfect score:  2908
Sequence:       1 atgagacgatccggaaacta..........tattcgagccctttgcatga 1632

Scoring table:  BLOSUM62
                Xgapop 10.0 , Xgapext  0.5
                Ygapop 10.0 , Ygapext  0.5
                Fgapop  6.0 , Fgapext  7.0
                Delop   6.0 , Delext   7.0

Searched:       389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters:      778828

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US09938956/runat_24022004_144702_3089/app_query.fasta_1.1799
-DB=Issued_Patents_AA -QFMT=fastan -SUFFIX=rai -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09938956_@CGN_1_1_27_@runat_24022004_144702_3089 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :      Issued_Patents_AA:*
                1:  /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
                2:  /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
                3:  /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
                4:  /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
                5:  /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
                6:  /cgn2_6/ptodata/2/iaa/backfiles1.pep:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2876	98.9	599	2	US-08-846-526-11	Sequence 11, Appl
2	2876	98.9	599	3	US-09-172-339-2	Sequence 2, Appli
3	2876	98.9	599	4	US-09-398-395A-22	Sequence 22, Appl
4	2876	98.9	599	4	US-09-887-586A-22	Sequence 22, Appl
5	2876	98.9	599	4	US-09-895-752-22	Sequence 22, Appl
6	2876	98.9	599	4	US-09-903-012B-22	Sequence 22, Appl
7	2876	98.9	599	4	US-09-900-797-22	Sequence 22, Appl
8	1633.5	56.2	591	2	US-08-937-540-4	Sequence 4, Appli
9	1633.5	56.2	591	4	US-09-398-395A-24	Sequence 24, Appl
10	1633.5	56.2	591	4	US-09-887-586A-24	Sequence 24, Appl
11	1633.5	56.2	591	4	US-09-895-752-24	Sequence 24, Appl
12	1633.5	56.2	591	4	US-09-903-012B-24	Sequence 24, Appl
13	1633.5	56.2	591	4	US-09-900-797-24	Sequence 24, Appl
14	1553.5	53.4	589	2	US-08-937-540-6	Sequence 6, Appli
15	1553.5	53.4	590	4	US-09-398-395A-54	Sequence 54, Appl
16	1553.5	53.4	590	4	US-09-887-586A-54	Sequence 54, Appl
17	1553.5	53.4	590	4	US-09-895-752-54	Sequence 54, Appl
18	1553.5	53.4	590	4	US-09-903-012B-54	Sequence 54, Appl
19	1553.5	53.4	590	4	US-09-900-797-54	Sequence 54, Appl
20	1468	50.5	598	2	US-08-937-540-2	Sequence 2, Appli
21	1468	50.5	598	4	US-09-398-395A-26	Sequence 26, Appl
22	1468	50.5	598	4	US-09-887-586A-26	Sequence 26, Appl
23	1468	50.5	598	4	US-09-895-752-26	Sequence 26, Appl
24	1468	50.5	598	4	US-09-903-012B-26	Sequence 26, Appl
25	1468	50.5	598	4	US-09-900-797-26	Sequence 26, Appl
26	1289.5	44.3	279	2	US-08-326-286-7	Sequence 7, Appli
27	836	28.7	556	4	US-09-398-395A-32	Sequence 32, Appl
28	836	28.7	556	4	US-09-887-586A-32	Sequence 32, Appl
29	836	28.7	556	4	US-09-895-752-32	Sequence 32, Appl
30	836	28.7	556	4	US-09-903-012B-32	Sequence 32, Appl
31	836	28.7	556	4	US-09-900-797-32	Sequence 32, Appl
32	809.5	27.8	550	2	US-08-443-639-8	Sequence 8, Appli
33	809.5	27.8	550	3	US-08-577-483-8	Sequence 8, Appli
34	809.5	27.8	550	4	US-09-435-380-8	Sequence 8, Appli
35	806.5	27.7	548	4	US-09-398-395A-2	Sequence 2, Appli
36	806.5	27.7	548	4	US-09-887-586A-2	Sequence 2, Appli
37	806.5	27.7	548	4	US-09-895-752-2	Sequence 2, Appli
38	806.5	27.7	548	4	US-09-903-012B-2	Sequence 2, Appli
39	806.5	27.7	548	4	US-09-900-797-2	Sequence 2, Appli
40	804.5	27.7	548	4	US-09-398-395A-8	Sequence 8, Appli
41	804.5	27.7	548	4	US-09-398-395A-12	Sequence 12, Appl
42	804.5	27.7	548	4	US-09-887-586A-8	Sequence 8, Appli
43	804.5	27.7	548	4	US-09-887-586A-12	Sequence 12, Appl
44	804.5	27.7	548	4	US-09-895-752-8	Sequence 8, Appli
45	804.5	27.7	548	4	US-09-895-752-12	Sequence 12, Appl

ALIGNMENTS

```
RESULT 1
US-08-846-526-11
; Sequence 11, Application US/08846526
; Patent No. 5871988
;  GENERAL INFORMATION:
;    APPLICANT:  Crouteau, Rodney B.
;    APPLICANT:  Colby, Sheila M.
;    TITLE OF INVENTION:  DNA Encoding Limonene Synthase From
;    TITLE OF INVENTION:  Mentha Spicata
;    NUMBER OF SEQUENCES:  11
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Christensen, O'Connor, Johnson & Kindness
;      STREET:  2800 Pacific First Centre, 1420 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  USA
;      ZIP:  98101-2347
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/846,526
;      FILING DATE:  29-APR-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/582,802
;      FILING DATE:
;      APPLICATION NUMBER:  US/08/145,941
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Shelton, Dennis K.
;      REGISTRATION NUMBER:  26,997
;      REFERENCE/DOCKET NUMBER:  WSUR-1-7219
;    TELECOMMUNICATION INFORMATION:
```

```
;       TELEPHONE:  (206) 224-0718
;       TELEFAX:  (206) 224-0779
;       TELEX:  4938023
;   INFORMATION FOR SEQ ID NO:  11:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  599 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US-08-846-526-11

Alignment Scores:
Pred. No.:               2.52e-284      Length:          599
Score:                   2876.00        Matches:         542
Percent Similarity:      100.00%        Conservative:    0
Best Local Similarity:   100.00%        Mismatches:      0
Query Match:             98.90%         Indels:          0
DB:                      2              Gaps:            0

US-09-938-956-6 (1-1632) x US-08-846-526-11 (1-599)

Qy        4 AGACGATCCGGAAACTACAACCCTTCTCGTTGGGATGTCAACTTCATCCAATCGCTTCTC 63
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       58 ArgArgSerGlyAsnTyrAsnProSerArgTrpAspValAsnPheIleGlnSerLeuLeu 77

Qy       64 AGTGACTATAAGGAGGACAAACACGTGATTAGGGCTTCTGAGCTGGTCACTTTGGTGAAG 123
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       78 SerAspTyrLysGluAspLysHisValIleArgAlaSerGluLeuValThrLeuValLys 97

Qy      124 ATGGAACTGGAGAAAGAAACGGATCAAATTCGACAACTTGAGTTGATCGATGACTTGCAG 183
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       98 MetGluLeuGluLysGluThrAspGlnIleArgGlnLeuGluLeuIleAspAspLeuGln 117

Qy      184 AGGATGGGGCTGTCCGATCATTTCCAAAATGAGTTCAAAGAAATCTTGTCCTCTATATAT 243
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      118 ArgMetGlyLeuSerAspHisPheGlnAsnGluPheLysGluIleLeuSerSerIleTyr 137

Qy      244 CTCGACCATCACTATTACAAGAACCCTTTTCCAAAAGAAGAAAGGGATCTCTACTCCACA 303
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      138 LeuAspHisHisTyrTyrLysAsnProPheProLysGluGluArgAspLeuTyrSerThr 157

Qy      304 TCTCTTGCATTTAGGCTCCTCAGAGAACATGGTTTTCAAGTCGCACAAGAGGTATTCGAT 363
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      158 SerLeuAlaPheArgLeuLeuArgGluHisGlyPheGlnValAlaGlnGluValPheAsp 177

Qy      364 AGTTTCAAGAACGAGGAGGGTGAGTTCAAAGAAAGCCTTAGCGACGACACCAGAGGATTG 423
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      178 SerPheLysAsnGluGluGlyGluPheLysGluSerLeuSerAspAspThrArgGlyLeu 197

Qy      424 TTGCAACTGTATGAAGCTTCCTTTCTGTTGACGGAAGGCGAAACCACGCTCGAGTCAGCG 483
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      198 LeuGlnLeuTyrGluAlaSerPheLeuLeuThrGluGlyGluThrThrLeuGluSerAla 217

Qy      484 AGGGAATTCGCCACCAAATTTTTGGAGGAAAAAGTGAACGAGGGTGGTGTTGATGGCGAC 543
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      218 ArgGluPheAlaThrLysPheLeuGluGluLysValAsnGluGlyGlyValAspGlyAsp 237

Qy      544 CTTTTAACAAGAATCGCATATTCTTTGGACATCCCTCTTCATTGGAGGATTAAAAGGCCA 603
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      238 LeuLeuThrArgIleAlaTyrSerLeuAspIleProLeuHisTrpArgIleLysArgPro 257

Qy      604 AATGCACCTGTGTGGATCGAATGGTATAGGAAGAGGCCCGACATGAATCCAGTAGTGTTG 663
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      258 AsnAlaProValTrpIleGluTrpTyrArgLysArgProAspMetAsnProValValLeu 277

Qy      664 GAGCTTGCCATACTCGACTTAAATATTGTTCAAGCACAATTTCAAGAAGAGCTCAAAGAA 723
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      278 GluLeuAlaIleLeuAspLeuAsnIleValGlnAlaGlnPheGlnGluGluLeuLysGlu 297

Qy      724 TCCTTCAGGTGGTGGAGAAATACTGGGTTTGTTGAGAAGCTGCCCTTCGCAAGGGATAGA 783
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      298 SerPheArgTrpTrpArgAsnThrGlyPheValGluLysLeuProPheAlaArgAspArg 317

Qy      784 CTGGTGGAATGCTACTTTTGGAATACTGGGATCATCGAGCCACGTCAGCATGCAAGTGCA 843
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      318 LeuValGluCysTyrPheTrpAsnThrGlyIleIleGluProArgGlnHisAlaSerAla 337

Qy      844 AGGATAATGATGGGCAAAGTCAACGCTCTGATTACGGTGATCGATGATATTTATGATGTC 903
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      338 ArgIleMetMetGlyLysValAsnAlaLeuIleThrValIleAspAspIleTyrAspVal 357

Qy      904 TATGGCACCTTAGAAGAACTCGAACAATTCACTGACCTCATTCGAAGATGGGATATAAAC 963
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      358 TyrGlyThrLeuGluGluLeuGluGlnPheThrAspLeuIleArgArgTrpAspIleAsn 377

Qy      964 TCAATCGACCAACTTCCCGATTACATGCAACTGTGCTTTCTTGCACTCAACAACTTCGTC 1023
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      378 SerIleAspGlnLeuProAspTyrMetGlnLeuCysPheLeuAlaLeuAsnAsnPheVal 397

Qy     1024 GATGATACATCGTACGATGTTATGAAGGAGAAAGGCGTCAACGTTATACCCTACCTGCGG 1083
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      398 AspAspThrSerTyrAspValMetLysGluLysGlyValAsnValIleProTyrLeuArg 417

Qy     1084 CAATCGTGGGTTGATTTGGCGGATAAGTATATGGTAGAGGCACGGTGGTTCTACGGCGGG 1143
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      418 GlnSerTrpValAspLeuAlaAspLysTyrMetValGluAlaArgTrpPheTyrGlyGly 437

Qy     1144 CACAAACCAAGTTTGGAAGAGTATTTGGAGAACTCATGGCAGTCGATAAGTGGGCCCTGT 1203
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      438 HisLysProSerLeuGluGluTyrLeuGluAsnSerTrpGlnSerIleSerGlyProCys 457

Qy     1204 ATGTTAACGCACATATTCTTCCGAGTAACAGATTCGTTCACAAAGGAGACCGTCGACAGT 1263
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      458 MetLeuThrHisIlePhePheArgValThrAspSerPheThrLysGluThrValAspSer 477

Qy     1264 TTGTACAAATACCACGATTTAGTTCGTTGGTCATCCTTCGTTCTGCGGCTTGCTGATGAT 1323
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      478 LeuTyrLysTyrHisAspLeuValArgTrpSerSerPheValLeuArgLeuAlaAspAsp 497

Qy     1324 TTGGGAACCTCGGTGGAAGAGGTGAGCAGAGGGGATGTGCCGAAATCACTTCAGTGCTAC 1383
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      498 LeuGlyThrSerValGluGluValSerArgGlyAspValProLysSerLeuGlnCysTyr 517

Qy     1384 ATGAGTGACTACAATGCATCGGAGGCGGAGGCGCGGAAGCACGTGAAATGGCTGATAGCG 1443
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      518 MetSerAspTyrAsnAlaSerGluAlaGluAlaArgLysHisValLysTrpLeuIleAla 537

Qy     1444 GAGGTGTGGAAGAAGATGAATGCGGAGAGGGTGTCGAAGGATTCTCCATTCGGCAAAGAT 1503
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      538 GluValTrpLysLysMetAsnAlaGluArgValSerLysAspSerProPheGlyLysAsp 557

Qy     1504 TTTATAGGATGTGCAGTTGATTTAGGAAGGATGGCGCAGTTGATGTACCATAATGGAGAT 1563
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      558 PheIleGlyCysAlaValAspLeuGlyArgMetAlaGlnLeuMetTyrHisAsnGlyAsp 577

Qy     1564 GGGCACGGCACACAACACCCTATTATACATCAACAAATGACCAGAACCTTATTCGAGCCC 1623
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      578 GlyHisGlyThrGlnHisProIleIleHisGlnGlnMetThrArgThrLeuPheGluPro 597

Qy     1624 TTTGCA 1629
            ||||||
Db      598 PheAla 599

RESULT 2
US-09-172-339-2
; Sequence 2, Application US/09172339
; Patent No. 6291745
; GENERAL INFORMATION:
;  APPLICANT: EuClaire Meyer, Terry
;  APPLICANT: Yalpani, Nasser
;  TITLE OF INVENTION: Limonene and Other Downstream
;  TITLE OF INVENTION: Metabolites of Geranyl Pyrophosphate for Insect Control in
;  TITLE OF INVENTION: Plants
;  FILE REFERENCE: 5718-65
;  CURRENT APPLICATION NUMBER: US/09/172,339
;  CURRENT FILING DATE: 1998-10-14
```

```
;  EARLIER APPLICATION NUMBER: 08/449,061
;  EARLIER FILING DATE: 1995-05-24
;  EARLIER APPLICATION NUMBER: 08/153,544
;  EARLIER FILING DATE: 1993-11-16
;  EARLIER APPLICATION NUMBER: 08/042,199
;  EARLIER FILING DATE: 1993-04-02
;  NUMBER OF SEQ ID NOS: 8
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2
;   LENGTH: 599
;   TYPE: PRT
;   ORGANISM: Mentha spicata
US-09-172-339-2

Alignment Scores:
Pred. No.:              2.52e-284      Length:         599
Score:                  2876.00        Matches:        542
Percent Similarity:     100.00%        Conservative:   0
Best Local Similarity:  100.00%        Mismatches:     0
Query Match:            98.90%         Indels:         0
DB:                     3              Gaps:           0

US-09-938-956-6 (1-1632) x US-09-172-339-2 (1-599)

Qy          4 AGACGATCCGGAAACTACAACCCTTCTCGTTGGGATGTCAACTTCATCCAATCGCTTCTC 63
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         58 ArgArgSerGlyAsnTyrAsnProSerArgTrpAspValAsnPheIleGlnSerLeuLeu 77

Qy         64 AGTGACTATAAGGAGGACAAACACGTGATTAGGGCTTCTGAGCTGGTCACTTTGGTGAAG 123
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         78 SerAspTyrLysGluAspLysHisValIleArgAlaSerGluLeuValThrLeuValLys 97

Qy        124 ATGGAACTGGAGAAAGAAACGGATCAAATTCGACAACTTGAGTTGATCGATGACTTGCAG 183
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         98 MetGluLeuGluLysGluThrAspGlnIleArgGlnLeuGluLeuIleAspAspLeuGln 117

Qy        184 AGGATGGGGCTGTCCGATCATTTCCAAAATGAGTTCAAAGAAATCTTGTCCTCTATATAT 243
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        118 ArgMetGlyLeuSerAspHisPheGlnAsnGluPheLysGluIleLeuSerSerIleTyr 137

Qy        244 CTCGACCATCACTATTACAAGAACCCTTTTCCAAAAGAAGAAAGGGATCTCTACTCCACA 303
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        138 LeuAspHisHisTyrTyrLysAsnProPheProLysGluGluArgAspLeuTyrSerThr 157

Qy        304 TCTCTTGCATTTAGGCTCCTCAGAGAACATGGTTTTCAAGTCGCACAAGAGGTATTCGAT 363
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        158 SerLeuAlaPheArgLeuLeuArgGluHisGlyPheGlnValAlaGlnGluValPheAsp 177

Qy        364 AGTTTCAAGAACGAGGAGGGTGAGTTCAAAGAAAGCCTTAGCGACGACACCAGAGGATTG 423
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        178 SerPheLysAsnGluGluGlyGluPheLysGluSerLeuSerAspAspThrArgGlyLeu 197

Qy        424 TTGCAACTGTATGAAGCTTCCTTTCTGTTGACGGAAGGCGAAACCACGCTCGAGTCAGCG 483
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        198 LeuGlnLeuTyrGluAlaSerPheLeuLeuThrGluGlyGluThrThrLeuGluSerAla 217

Qy        484 AGGGAATTCGCCACCAAATTTTTGGAGGAAAAAGTGAACGAGGGTGGTGTTGATGGCGAC 543
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        218 ArgGluPheAlaThrLysPheLeuGluGluLysValAsnGluGlyGlyValAspGlyAsp 237

Qy        544 CTTTTAACAAGAATCGCATATTCTTTGGACATCCCTCTTCATTGGAGGATTAAAAGGCCA 603
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        238 LeuLeuThrArgIleAlaTyrSerLeuAspIleProLeuHisTrpArgIleLysArgPro 257

Qy        604 AATGCACCTGTGTGGATCGAATGGTATAGGAAGAGGCCCGACATGAATCCAGTAGTGTTG 663
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        258 AsnAlaProValTrpIleGluTrpTyrArgLysArgProAspMetAsnProValValLeu 277

Qy        664 GAGCTTGCCATACTCGACTTAAATATTGTTCAAGCACAATTTCAAGAAGAGCTCAAAGAA 723
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        278 GluLeuAlaIleLeuAspLeuAsnIleValGlnAlaGlnPheGlnGluGluLeuLysGlu 297

Qy        724 TCCTTCAGGTGGTGGAGAAATACTGGGTTTGTTGAGAAGCTGCCCTTCGCAAGGGATAGA 783
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        298 SerPheArgTrpTrpArgAsnThrGlyPheValGluLysLeuProPheAlaArgAspArg 317

Qy        784 CTGGTGGAATGCTACTTTTGGAATACTGGGATCATCGAGCCACGTCAGCATGCAAGTGCA 843
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        318 LeuValGluCysTyrPheTrpAsnThrGlyIleIleGluProArgGlnHisAlaSerAla 337

Qy        844 AGGATAATGATGGGCAAAGTCAACGCTCTGATTACGGTGATCGATGATATTTATGATGTC 903
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        338 ArgIleMetMetGlyLysValAsnAlaLeuIleThrValIleAspAspIleTyrAspVal 357

Qy        904 TATGGCACCTTAGAAGAACTCGAACAATTCACTGACCTCATTCGAAGATGGGATATAAAC 963
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        358 TyrGlyThrLeuGluGluLeuGluGlnPheThrAspLeuIleArgArgTrpAspIleAsn 377

Qy        964 TCAATCGACCAACTTCCCGATTACATGCAACTGTGCTTTCTTGCACTCAACAACTTCGTC 1023
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        378 SerIleAspGlnLeuProAspTyrMetGlnLeuCysPheLeuAlaLeuAsnAsnPheVal 397

Qy       1024 GATGATACATCGTACGATGTTATGAAGGAGAAAGGCGTCAACGTTATACCCTACCTGCGG 1083
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        398 AspAspThrSerTyrAspValMetLysGluLysGlyValAsnValIleProTyrLeuArg 417

Qy       1084 CAATCGTGGGTTGATTTGGCGGATAAGTATATGGTAGAGGCACGGTGGTTCTACGGCGGG 1143
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        418 GlnSerTrpValAspLeuAlaAspLysTyrMetValGluAlaArgTrpPheTyrGlyGly 437

Qy       1144 CACAAACCAAGTTTGGAAGAGTATTTGGAGAACTCATGGCAGTCGATAAGTGGGCCCTGT 1203
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        438 HisLysProSerLeuGluGluTyrLeuGluAsnSerTrpGlnSerIleSerGlyProCys 457

Qy       1204 ATGTTAACGCACATATTCTTCCGAGTAACAGATTCGTTCACAAAGGAGACCGTCGACAGT 1263
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        458 MetLeuThrHisIlePhePheArgValThrAspSerPheThrLysGluThrValAspSer 477

Qy       1264 TTGTACAAATACCACGATTTAGTTCGTTGGTCATCCTTCGTTCTGCGGCTTGCTGATGAT 1323
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        478 LeuTyrLysTyrHisAspLeuValArgTrpSerSerPheValLeuArgLeuAlaAspAsp 497

Qy       1324 TTGGGAACCTCGGTGGAAGAGGTGAGCAGAGGGGATGTGCCGAAATCACTTCAGTGCTAC 1383
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        498 LeuGlyThrSerValGluGluValSerArgGlyAspValProLysSerLeuGlnCysTyr 517

Qy       1384 ATGAGTGACTACAATGCATCGGAGGCGGAGGCGCGGAAGCACGTGAAATGGCTGATAGCG 1443
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        518 MetSerAspTyrAsnAlaSerGluAlaGluAlaArgLysHisValLysTrpLeuIleAla 537

Qy       1444 GAGGTGTGGAAGAAGATGAATGCGGAGAGGGTGTCGAAGGATTCTCCATTCGGCAAAGAT 1503
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        538 GluValTrpLysLysMetAsnAlaGluArgValSerLysAspSerProPheGlyLysAsp 557

Qy       1504 TTTATAGGATGTGCAGTTGATTTAGGAAGGATGGCGCAGTTGATGTACCATAATGGAGAT 1563
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        558 PheIleGlyCysAlaValAspLeuGlyArgMetAlaGlnLeuMetTyrHisAsnGlyAsp 577

Qy       1564 GGGCACGGCACACAACACCCTATTATACATCAACAAATGACCAGAACCTTATTCGAGCCC 1623
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        578 GlyHisGlyThrGlnHisProIleIleHisGlnGlnMetThrArgThrLeuPheGluPro 597

Qy       1624 TTTGCA 1629
              ||||||
Db        598 PheAla 599

RESULT 3
US-09-398-395A-22
; Sequence 22, Application US/09398395A
; Patent No. 6468772
; GENERAL INFORMATION:
;  APPLICANT: Chappell, Joseph
;  APPLICANT:  No. 64687721, Joseph P.
;  APPLICANT:  Starks, Courtney M.
;  APPLICANT:  Manna, Kathleen R.
;  TITLE OF INVENTION: SYNTHASES
```

```
;  FILE REFERENCE: 07678-025001
;  CURRENT APPLICATION NUMBER: US/09/398,395A
;  CURRENT FILING DATE:  1999-09-17
;  PRIOR APPLICATION NUMBER: 60/100,993
;  PRIOR FILING DATE: 1998-09-18
;  PRIOR APPLICATION NUMBER: 60/130,628
;  PRIOR FILING DATE: 1999-04-22
;  PRIOR APPLICATION NUMBER: 60/150,262
;  PRIOR FILING DATE: 1999-08-23
;  NUMBER OF SEQ ID NOS: 58
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 22
;   LENGTH: 599
;   TYPE: PRT
;   ORGANISM: Mentha spicata
US-09-398-395A-22

Alignment Scores:
Pred. No.:              2.52e-284     Length:         599
Score:                  2876.00       Matches:        542
Percent Similarity:     100.00%       Conservative:   0
Best Local Similarity:  100.00%       Mismatches:     0
Query Match:            98.90%        Indels:         0
DB:                     4             Gaps:           0

US-09-938-956-6 (1-1632) x US-09-398-395A-22 (1-599)

Qy        4 AGACGATCCGGAAACTACAACCCTTCTCGTTGGGATGTCAACTTCATCCAATCGCTTCTC 63
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       58 ArgArgSerGlyAsnTyrAsnProSerArgTrpAspValAsnPheIleGlnSerLeuLeu 77

Qy       64 AGTGACTATAAGGAGGACAAACACGTGATTAGGGCTTCTGAGCTGGTCACTTTGGTGAAG 123
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       78 SerAspTyrLysGluAspLysHisValIleArgAlaSerGluLeuValThrLeuValLys 97

Qy      124 ATGGAACTGGAGAAAGAAACGGATCAAATTCGACAACTTGAGTTGATCGATGACTTGCAG 183
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       98 MetGluLeuGluLysGluThrAspGlnIleArgGlnLeuGluLeuIleAspAspLeuGln 117

Qy      184 AGGATGGGGCTGTCCGATCATTTCCAAAATGAGTTCAAAGAAATCTTGTCCTCTATATAT 243
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      118 ArgMetGlyLeuSerAspHisPheGlnAsnGluPheLysGluIleLeuSerSerIleTyr 137

Qy      244 CTCGACCATCACTATTACAAGAACCCTTTTCCAAAAGAAGAAAGGGATCTCTACTCCACA 303
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      138 LeuAspHisHisTyrTyrLysAsnProPheProLysGluGluArgAspLeuTyrSerThr 157

Qy      304 TCTCTTGCATTTAGGCTCCTCAGAGAACATGGTTTTCAAGTCGCACAAGAGGTATTCGAT 363
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      158 SerLeuAlaPheArgLeuLeuArgGluHisGlyPheGlnValAlaGlnGluValPheAsp 177

Qy      364 AGTTTCAAGAACGAGGAGGGTGAGTTCAAAGAAAGCCTTAGCGACGACACCAGAGGATTG 423
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      178 SerPheLysAsnGluGluGlyGluPheLysGluSerLeuSerAspAspThrArgGlyLeu 197

Qy      424 TTGCAACTGTATGAAGCTTCCTTTCTGTTGACGGAAGGCGAAACCACGCTCGAGTCAGCG 483
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      198 LeuGlnLeuTyrGluAlaSerPheLeuLeuThrGluGlyGluThrThrLeuGluSerAla 217

Qy      484 AGGGAATTCGCCACCAAATTTTTGGAGGAAAAAGTGAACGAGGGTGGTGTTGATGGCGAC 543
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      218 ArgGluPheAlaThrLysPheLeuGluGluLysValAsnGluGlyGlyValAspGlyAsp 237

Qy      544 CTTTTAACAAGAATCGCATATTCTTTGGACATCCCTCTTCATTGGAGGATTAAAAGGCCA 603
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      238 LeuLeuThrArgIleAlaTyrSerLeuAspIleProLeuHisTrpArgIleLysArgPro 257

Qy      604 AATGCACCTGTGTGGATCGAATGGTATAGGAAGAGGCCCGACATGAATCCAGTAGTGTTG 663
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      258 AsnAlaProValTrpIleGluTrpTyrArgLysArgProAspMetAsnProValValLeu 277

Qy      664 GAGCTTGCCATACTCGACTTAAATATTGTTCAAGCACAATTTCAAGAAGAGCTCAAAGAA 723
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      278 GluLeuAlaIleLeuAspLeuAsnIleValGlnAlaGlnPheGlnGluGluLeuLysGlu 297

Qy      724 TCCTTCAGGTGGTGGAGAAATACTGGGTTTGTTGAGAAGCTGCCCTTCGCAAGGGATAGA 783
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      298 SerPheArgTrpTrpArgAsnThrGlyPheValGluLysLeuProPheAlaArgAspArg 317

Qy      784 CTGGTGGAATGCTACTTTTGGAATACTGGGATCATCGAGCCACGTCAGCATGCAAGTGCA 843
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      318 LeuValGluCysTyrPheTrpAsnThrGlyIleIleGluProArgGlnHisAlaSerAla 337

Qy      844 AGGATAATGATGGGCAAAGTCAACGCTCTGATTACGGTGATCGATGATATTTATGATGTC 903
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      338 ArgIleMetMetGlyLysValAsnAlaLeuIleThrValIleAspAspIleTyrAspVal 357

Qy      904 TATGGCACCTTAGAAGAACTCGAACAATTCACTGACCTCATTCGAAGATGGGATATAAAC 963
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      358 TyrGlyThrLeuGluGluLeuGluGlnPheThrAspLeuIleArgArgTrpAspIleAsn 377

Qy      964 TCAATCGACCAACTTCCCGATTACATGCAACTGTGCTTTCTTGCACTCAACAACTTCGTC 1023
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      378 SerIleAspGlnLeuProAspTyrMetGlnLeuCysPheLeuAlaLeuAsnAsnPheVal 397

Qy     1024 GATGATACATCGTACGATGTTATGAAGGAGAAAGGCGTCAACGTTATACCCTACCTGCGG 1083
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      398 AspAspThrSerTyrAspValMetLysGluLysGlyValAsnValIleProTyrLeuArg 417

Qy     1084 CAATCGTGGGTTGATTTGGCGGATAAGTATATGGTAGAGGCACGGTGGTTCTACGGCGGG 1143
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      418 GlnSerTrpValAspLeuAlaAspLysTyrMetValGluAlaArgTrpPheTyrGlyGly 437

Qy     1144 CACAAACCAAGTTTGGAAGAGTATTTGGAGAACTCATGGCAGTCGATAAGTGGGCCCTGT 1203
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      438 HisLysProSerLeuGluGluTyrLeuGluAsnSerTrpGlnSerIleSerGlyProCys 457

Qy     1204 ATGTTAACGCACATATTCTTCCGAGTAACAGATTCGTTCACAAAGGAGACCGTCGACAGT 1263
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      458 MetLeuThrHisIlePhePheArgValThrAspSerPheThrLysGluThrValAspSer 477

Qy     1264 TTGTACAAATACCACGATTTAGTTCGTTGGTCATCCTTCGTTCTGCGGCTTGCTGATGAT 1323
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      478 LeuTyrLysTyrHisAspLeuValArgTrpSerSerPheValLeuArgLeuAlaAspAsp 497

Qy     1324 TTGGGAACCTCGGTGGAAGAGGTGAGCAGAGGGGATGTGCCGAAATCACTTCAGTGCTAC 1383
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      498 LeuGlyThrSerValGluGluValSerArgGlyAspValProLysSerLeuGlnCysTyr 517

Qy     1384 ATGAGTGACTACAATGCATCGGAGGCGGAGGCGCGGAAGCACGTGAAATGGCTGATAGCG 1443
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      518 MetSerAspTyrAsnAlaSerGluAlaGluAlaArgLysHisValLysTrpLeuIleAla 537

Qy     1444 GAGGTGTGGAAGAAGATGAATGCGGAGAGGGTGTCGAAGGATTCTCCATTCGGCAAAGAT 1503
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      538 GluValTrpLysLysMetAsnAlaGluArgValSerLysAspSerProPheGlyLysAsp 557

Qy     1504 TTTATAGGATGTGCAGTTGATTTAGGAAGGATGGCGCAGTTGATGTACCATAATGGAGAT 1563
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      558 PheIleGlyCysAlaValAspLeuGlyArgMetAlaGlnLeuMetTyrHisAsnGlyAsp 577

Qy     1564 GGGCACGGCACACAACACCCTATTATACATCAACAAATGACCAGAACCTTATTCGAGCCC 1623
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      578 GlyHisGlyThrGlnHisProIleIleHisGlnGlnMetThrArgThrLeuPheGluPro 597

Qy     1624 TTTGCA 1629
            ||||||
Db      598 PheAla 599

RESULT 4
US-09-887-586A-22
; Sequence 22, Application US/09887586A
; Patent No. 6495354
; GENERAL INFORMATION:
;  APPLICANT: Chappell, Joseph
;  APPLICANT:  No. 64953541, Joseph P.
```

```
;  APPLICANT:  Starks, Courtney M.
;  APPLICANT:  Manna, Kathleen R.
;  TITLE OF INVENTION: SYNTHASES
;  FILE REFERENCE: 07678-025001
;  CURRENT APPLICATION NUMBER: US/09/887,586A
;  CURRENT FILING DATE:  2001-06-22
;  PRIOR APPLICATION NUMBER: 09/398,395
;  PRIOR FILING DATE: 1999-09-17
;  PRIOR APPLICATION NUMBER: 60/130,628
;  PRIOR FILING DATE: 1999-04-22
;  PRIOR APPLICATION NUMBER: 60/150,262
;  PRIOR FILING DATE: 1999-08-23
;  NUMBER OF SEQ ID NOS: 58
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 22
;   LENGTH: 599
;   TYPE: PRT
;   ORGANISM: Mentha spicata
US-09-887-586A-22

Alignment Scores:
Pred. No.:              2.52e-284      Length:         599
Score:                  2876.00        Matches:        542
Percent Similarity:     100.00%        Conservative:   0
Best Local Similarity:  100.00%        Mismatches:     0
Query Match:            98.90%         Indels:         0
DB:                     4              Gaps:           0

US-09-938-956-6 (1-1632) x US-09-887-586A-22 (1-599)

Qy          4 AGACGATCCGGAAACTACAACCCTTCTCGTTGGGATGTCAACTTCATCCAATCGCTTCTC 63
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         58 ArgArgSerGlyAsnTyrAsnProSerArgTrpAspValAsnPheIleGlnSerLeuLeu 77

Qy         64 AGTGACTATAAGGAGGACAAACACGTGATTAGGGCTTCTGAGCTGGTCACTTTGGTGAAG 123
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         78 SerAspTyrLysGluAspLysHisValIleArgAlaSerGluLeuValThrLeuValLys 97

Qy        124 ATGGAACTGGAGAAAGAAACGGATCAAATTCGACAACTTGAGTTGATCGATGACTTGCAG 183
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         98 MetGluLeuGluLysGluThrAspGlnIleArgGlnLeuGluLeuIleAspAspLeuGln 117

Qy        184 AGGATGGGGCTGTCCGATCATTTCCAAAATGAGTTCAAAGAAATCTTGTCCTCTATATAT 243
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        118 ArgMetGlyLeuSerAspHisPheGlnAsnGluPheLysGluIleLeuSerSerIleTyr 137

Qy        244 CTCGACCATCACTATTACAAGAACCCTTTTCCAAAAGAAGAAAGGGATCTCTACTCCACA 303
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        138 LeuAspHisHisTyrTyrLysAsnProPheProLysGluGluArgAspLeuTyrSerThr 157

Qy        304 TCTCTTGCATTTAGGCTCCTCAGAGAACATGGTTTTCAAGTCGCACAAGAGGTATTCGAT 363
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        158 SerLeuAlaPheArgLeuLeuArgGluHisGlyPheGlnValAlaGlnGluValPheAsp 177

Qy        364 AGTTTCAAGAACGAGGAGGGTGAGTTCAAAGAAAGCCTTAGCGACGACACCAGAGGATTG 423
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        178 SerPheLysAsnGluGluGlyGluPheLysGluSerLeuSerAspAspThrArgGlyLeu 197

Qy        424 TTGCAACTGTATGAAGCTTCCTTTCTGTTGACGGAAGGCGAAACCACGCTCGAGTCAGCG 483
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        198 LeuGlnLeuTyrGluAlaSerPheLeuLeuThrGluGlyGluThrThrLeuGluSerAla 217

Qy        484 AGGGAATTCGCCACCAAATTTTTGGAGGAAAAAGTGAACGAGGGTGGTGTTGATGGCGAC 543
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        218 ArgGluPheAlaThrLysPheLeuGluGluLysValAsnGluGlyGlyValAspGlyAsp 237

Qy        544 CTTTTAACAAGAATCGCATATTCTTTGGACATCCCTCTTCATTGGAGGATTAAAAGGCCA 603
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        238 LeuLeuThrArgIleAlaTyrSerLeuAspIleProLeuHisTrpArgIleLysArgPro 257

Qy        604 AATGCACCTGTGTGGATCGAATGGTATAGGAAGAGGCCCGACATGAATCCAGTAGTGTTG 663
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        258 AsnAlaProValTrpIleGluTrpTyrArgLysArgProAspMetAsnProValValLeu 277

Qy        664 GAGCTTGCCATACTCGACTTAAATATTGTTCAAGCACAATTTCAAGAAGAGCTCAAAGAA 723
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        278 GluLeuAlaIleLeuAspLeuAsnIleValGlnAlaGlnPheGlnGluGluLeuLysGlu 297

Qy        724 TCCTTCAGGTGGTGGAGAAATACTGGGTTTGTTGAGAAGCTGCCCTTCGCAAGGGATAGA 783
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        298 SerPheArgTrpTrpArgAsnThrGlyPheValGluLysLeuProPheAlaArgAspArg 317

Qy        784 CTGGTGGAATGCTACTTTTGGAATACTGGGATCATCGAGCCACGTCAGCATGCAAGTGCA 843
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        318 LeuValGluCysTyrPheTrpAsnThrGlyIleIleGluProArgGlnHisAlaSerAla 337

Qy        844 AGGATAATGATGGGCAAAGTCAACGCTCTGATTACGGTGATCGATGATATTTATGATGTC 903
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        338 ArgIleMetMetGlyLysValAsnAlaLeuIleThrValIleAspAspIleTyrAspVal 357

Qy        904 TATGGCACCTTAGAAGAACTCGAACAATTCACTGACCTCATTCGAAGATGGGATATAAAC 963
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        358 TyrGlyThrLeuGluGluLeuGluGlnPheThrAspLeuIleArgArgTrpAspIleAsn 377

Qy        964 TCAATCGACCAACTTCCCGATTACATGCAACTGTGCTTTCTTGCACTCAACAACTTCGTC 1023
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        378 SerIleAspGlnLeuProAspTyrMetGlnLeuCysPheLeuAlaLeuAsnAsnPheVal 397

Qy       1024 GATGATACATCGTACGATGTTATGAAGGAGAAAGGCGTCAACGTTATACCCTACCTGCGG 1083
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        398 AspAspThrSerTyrAspValMetLysGluLysGlyValAsnValIleProTyrLeuArg 417

Qy       1084 CAATCGTGGGTTGATTTGGCGGATAAGTATATGGTAGAGGCACGGTGGTTCTACGGCGGG 1143
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        418 GlnSerTrpValAspLeuAlaAspLysTyrMetValGluAlaArgTrpPheTyrGlyGly 437

Qy       1144 CACAAACCAAGTTTGGAAGAGTATTTGGAGAACTCATGGCAGTCGATAAGTGGGCCCTGT 1203
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        438 HisLysProSerLeuGluGluTyrLeuGluAsnSerTrpGlnSerIleSerGlyProCys 457

Qy       1204 ATGTTAACGCACATATTCTTCCGAGTAACAGATTCGTTCACAAAGGAGACCGTCGACAGT 1263
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        458 MetLeuThrHisIlePhePheArgValThrAspSerPheThrLysGluThrValAspSer 477

Qy       1264 TTGTACAAATACCACGATTTAGTTCGTTGGTCATCCTTCGTTCTGCGGCTTGCTGATGAT 1323
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        478 LeuTyrLysTyrHisAspLeuValArgTrpSerSerPheValLeuArgLeuAlaAspAsp 497

Qy       1324 TTGGGAACCTCGGTGGAAGAGGTGAGCAGAGGGGATGTGCCGAAATCACTTCAGTGCTAC 1383
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        498 LeuGlyThrSerValGluGluValSerArgGlyAspValProLysSerLeuGlnCysTyr 517

Qy       1384 ATGAGTGACTACAATGCATCGGAGGCGGAGGCGCGGAAGCACGTGAAATGGCTGATAGCG 1443
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        518 MetSerAspTyrAsnAlaSerGluAlaGluAlaArgLysHisValLysTrpLeuIleAla 537

Qy       1444 GAGGTGTGGAAGAAGATGAATGCGGAGAGGGTGTCGAAGGATTCTCCATTCGGCAAAGAT 1503
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        538 GluValTrpLysLysMetAsnAlaGluArgValSerLysAspSerProPheGlyLysAsp 557

Qy       1504 TTTATAGGATGTGCAGTTGATTTAGGAAGGATGGCGCAGTTGATGTACCATAATGGAGAT 1563
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        558 PheIleGlyCysAlaValAspLeuGlyArgMetAlaGlnLeuMetTyrHisAsnGlyAsp 577

Qy       1564 GGGCACGGCACACAACACCCTATTATACATCAACAAATGACCAGAACCTTATTCGAGCCC 1623
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        578 GlyHisGlyThrGlnHisProIleIleHisGlnGlnMetThrArgThrLeuPheGluPro 597

Qy       1624 TTTGCA 1629
              ||||||
Db        598 PheAla 599

RESULT 5
US-09-895-752-22
; Sequence 22, Application US/09895752
; Patent No. 6559297
```

```
; GENERAL INFORMATION:
;  APPLICANT: Chappell, Joseph
;  APPLICANT:  No. 65592971, Joseph P.
;  APPLICANT:  Starks, Courtney M.
;  APPLICANT:  Manna, Kathleen R.
;  TITLE OF INVENTION: SYNTHASES
;  FILE REFERENCE: 07678-025001
;  CURRENT APPLICATION NUMBER: US/09/895,752
;  CURRENT FILING DATE:  2001-06-29
;  PRIOR APPLICATION NUMBER: 09/398,395
;  PRIOR FILING DATE: 1999-09-17
;  PRIOR APPLICATION NUMBER: 60/100,993
;  PRIOR FILING DATE: 1998-09-18
;  PRIOR APPLICATION NUMBER: 60/130,628
;  PRIOR FILING DATE: 1999-04-22
;  PRIOR APPLICATION NUMBER: 60/150,262
;  PRIOR FILING DATE: 1999-08-23
;  NUMBER OF SEQ ID NOS: 58
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 22
;   LENGTH: 599
;   TYPE: PRT
;   ORGANISM: Mentha spicata
US-09-895-752-22

Alignment Scores:
Pred. No.:               2.52e-284     Length:         599
Score:                   2876.00       Matches:        542
Percent Similarity:      100.00%       Conservative:   0
Best Local Similarity:   100.00%       Mismatches:     0
Query Match:             98.90%        Indels:         0
DB:                      4             Gaps:           0

US-09-938-956-6 (1-1632) x US-09-895-752-22 (1-599)

Qy          4 AGACGATCCGGAAACTACAACCCTTCTCGTTGGGATGTCAACTTCATCCAATCGCTTCTC 63
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         58 ArgArgSerGlyAsnTyrAsnProSerArgTrpAspValAsnPheIleGlnSerLeuLeu 77

Qy         64 AGTGACTATAAGGAGGACAAACACGTGATTAGGGCTTCTGAGCTGGTCACTTTGGTGAAG 123
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         78 SerAspTyrLysGluAspLysHisValIleArgAlaSerGluLeuValThrLeuValLys 97

Qy        124 ATGGAACTGGAGAAAGAAACGGATCAAATTCGACAACTTGAGTTGATCGATGACTTGCAG 183
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         98 MetGluLeuGluLysGluThrAspGlnIleArgGlnLeuGluLeuIleAspAspLeuGln 117

Qy        184 AGGATGGGGCTGTCCGATCATTTCCAAAATGAGTTCAAAGAAATCTTGTCCTCTATATAT 243
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        118 ArgMetGlyLeuSerAspHisPheGlnAsnGluPheLysGluIleLeuSerSerIleTyr 137

Qy        244 CTCGACCATCACTATTACAAGAACCCTTTTCCAAAAGAAGAAAGGGATCTCTACTCCACA 303
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        138 LeuAspHisHisTyrTyrLysAsnProPheProLysGluGluArgAspLeuTyrSerThr 157

Qy        304 TCTCTTGCATTTAGGCTCCTCAGAGAACATGGTTTTCAAGTCGCACAAGAGGTATTCGAT 363
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        158 SerLeuAlaPheArgLeuLeuArgGluHisGlyPheGlnValAlaGlnGluValPheAsp 177

Qy        364 AGTTTCAAGAACGAGGAGGGTGAGTTCAAAGAAAGCCTTAGCGACGACACCAGAGGATTG 423
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        178 SerPheLysAsnGluGluGlyGluPheLysGluSerLeuSerAspAspThrArgGlyLeu 197

Qy        424 TTGCAACTGTATGAAGCTTCCTTTCTGTTGACGGAAGGCGAAACCACGCTCGAGTCAGCG 483
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        198 LeuGlnLeuTyrGluAlaSerPheLeuLeuThrGluGlyGluThrThrLeuGluSerAla 217

Qy        484 AGGGAATTCGCCACCAAATTTTTGGAGGAAAAAGTGAACGAGGGTGGTGTTGATGGCGAC 543
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        218 ArgGluPheAlaThrLysPheLeuGluGluLysValAsnGluGlyGlyValAspGlyAsp 237

Qy        544 CTTTTAACAAGAATCGCATATTCTTTGGACATCCCTCTTCATTGGAGGATTAAAAGGCCA 603
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        238 LeuLeuThrArgIleAlaTyrSerLeuAspIleProLeuHisTrpArgIleLysArgPro 257

Qy        604 AATGCACCTGTGTGGATCGAATGGTATAGGAAGAGGCCCGACATGAATCCAGTAGTGTTG 663
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        258 AsnAlaProValTrpIleGluTrpTyrArgLysArgProAspMetAsnProValValLeu 277

Qy        664 GAGCTTGCCATACTCGACTTAAATATTGTTCAAGCACAATTTCAAGAAGAGCTCAAAGAA 723
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        278 GluLeuAlaIleLeuAspLeuAsnIleValGlnAlaGlnPheGlnGluGluLeuLysGlu 297

Qy        724 TCCTTCAGGTGGTGGAGAAATACTGGGTTTGTTGAGAAGCTGCCCTTCGCAAGGGATAGA 783
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        298 SerPheArgTrpTrpArgAsnThrGlyPheValGluLysLeuProPheAlaArgAspArg 317

Qy        784 CTGGTGGAATGCTACTTTTGGAATACTGGGATCATCGAGCCACGTCAGCATGCAAGTGCA 843
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        318 LeuValGluCysTyrPheTrpAsnThrGlyIleIleGluProArgGlnHisAlaSerAla 337

Qy        844 AGGATAATGATGGGCAAAGTCAACGCTCTGATTACGGTGATCGATGATATTTATGATGTC 903
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        338 ArgIleMetMetGlyLysValAsnAlaLeuIleThrValIleAspAspIleTyrAspVal 357

Qy        904 TATGGCACCTTAGAAGAACTCGAACAATTCACTGACCTCATTCGAAGATGGGATATAAAC 963
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        358 TyrGlyThrLeuGluGluLeuGluGlnPheThrAspLeuIleArgArgTrpAspIleAsn 377

Qy        964 TCAATCGACCAACTTCCCGATTACATGCAACTGTGCTTTCTTGCACTCAACAACTTCGTC 1023
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        378 SerIleAspGlnLeuProAspTyrMetGlnLeuCysPheLeuAlaLeuAsnAsnPheVal 397

Qy       1024 GATGATACATCGTACGATGTTATGAAGGAGAAAGGCGTCAACGTTATACCCTACCTGCGG 1083
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        398 AspAspThrSerTyrAspValMetLysGluLysGlyValAsnValIleProTyrLeuArg 417

Qy       1084 CAATCGTGGGTTGATTTGGCGGATAAGTATATGGTAGAGGCACGGTGGTTCTACGGCGGG 1143
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        418 GlnSerTrpValAspLeuAlaAspLysTyrMetValGluAlaArgTrpPheTyrGlyGly 437

Qy       1144 CACAAACCAAGTTTGGAAGAGTATTTGGAGAACTCATGGCAGTCGATAAGTGGGCCCTGT 1203
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        438 HisLysProSerLeuGluGluTyrLeuGluAsnSerTrpGlnSerIleSerGlyProCys 457

Qy       1204 ATGTTAACGCACATATTCTTCCGAGTAACAGATTCGTTCACAAAGGAGACCGTCGACAGT 1263
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        458 MetLeuThrHisIlePhePheArgValThrAspSerPheThrLysGluThrValAspSer 477

Qy       1264 TTGTACAAATACCACGATTTAGTTCGTTGGTCATCCTTCGTTCTGCGGCTTGCTGATGAT 1323
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        478 LeuTyrLysTyrHisAspLeuValArgTrpSerSerPheValLeuArgLeuAlaAspAsp 497

Qy       1324 TTGGGAACCTCGGTGGAAGAGGTGAGCAGAGGGGATGTGCCGAAATCACTTCAGTGCTAC 1383
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        498 LeuGlyThrSerValGluGluValSerArgGlyAspValProLysSerLeuGlnCysTyr 517

Qy       1384 ATGAGTGACTACAATGCATCGGAGGCGGAGGCGCGGAAGCACGTGAAATGGCTGATAGCG 1443
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        518 MetSerAspTyrAsnAlaSerGluAlaGluAlaArgLysHisValLysTrpLeuIleAla 537

Qy       1444 GAGGTGTGGAAGAAGATGAATGCGGAGAGGGTGTCGAAGGATTCTCCATTCGGCAAAGAT 1503
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        538 GluValTrpLysLysMetAsnAlaGluArgValSerLysAspSerProPheGlyLysAsp 557

Qy       1504 TTTATAGGATGTGCAGTTGATTTAGGAAGGATGGCGCAGTTGATGTACCATAATGGAGAT 1563
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        558 PheIleGlyCysAlaValAspLeuGlyArgMetAlaGlnLeuMetTyrHisAsnGlyAsp 577

Qy       1564 GGGCACGGCACACAACACCCTATTATACATCAACAAATGACCAGAACCTTATTCGAGCCC 1623
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        578 GlyHisGlyThrGlnHisProIleIleHisGlnGlnMetThrArgThrLeuPheGluPro 597

Qy       1624 TTTGCA 1629
              ||||||
Db        598 PheAla 599
```

```
RESULT 6
US-09-903-012B-22
; Sequence 22, Application US/09903012B
; Patent No. 6569656
; GENERAL INFORMATION:
;  APPLICANT: Chappell, Joseph
;  APPLICANT:  No. 65696561, Joseph P.
;  APPLICANT:  Starks, Courtney M.
;  APPLICANT:  Manna, Kathleen R.
;  TITLE OF INVENTION: SYNTHASES
;  FILE REFERENCE: 07678-025001
;  CURRENT APPLICATION NUMBER: US/09/903,012B
;  CURRENT FILING DATE:  2001-07-11
;  PRIOR APPLICATION NUMBER: 60/100,993
;  PRIOR FILING DATE: 1998-09-18
;  PRIOR APPLICATION NUMBER: 60/130,628
;  PRIOR FILING DATE: 1999-04-22
;  PRIOR APPLICATION NUMBER: 60/150,262
;  PRIOR FILING DATE: 1999-08-23
;  NUMBER OF SEQ ID NOS: 58
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 22
;   LENGTH: 599
;   TYPE: PRT
;   ORGANISM: Mentha spicata
US-09-903-012B-22

Alignment Scores:
Pred. No.:              2.52e-284     Length:         599
Score:                  2876.00       Matches:        542
Percent Similarity:     100.00%       Conservative:   0
Best Local Similarity:  100.00%       Mismatches:     0
Query Match:            98.90%        Indels:         0
DB:                     4             Gaps:           0

US-09-938-956-6 (1-1632) x US-09-903-012B-22 (1-599)

Qy        4 AGACGATCCGGAAACTACAACCCTTCTCGTTGGGATGTCAACTTCATCCAATCGCTTCTC 63
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       58 ArgArgSerGlyAsnTyrAsnProSerArgTrpAspValAsnPheIleGlnSerLeuLeu 77

Qy       64 AGTGACTATAAGGAGGACAAACACGTGATTAGGGCTTCTGAGCTGGTCACTTTGGTGAAG 123
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       78 SerAspTyrLysGluAspLysHisValIleArgAlaSerGluLeuValThrLeuValLys 97

Qy      124 ATGGAACTGGAGAAAGAAACGGATCAAATTCGACAACTTGAGTTGATCGATGACTTGCAG 183
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       98 MetGluLeuGluLysGluThrAspGlnIleArgGlnLeuGluLeuIleAspAspLeuGln 117

Qy      184 AGGATGGGGCTGTCCGATCATTTCCAAAATGAGTTCAAAGAAATCTTGTCCTCTATATAT 243
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      118 ArgMetGlyLeuSerAspHisPheGlnAsnGluPheLysGluIleLeuSerSerIleTyr 137

Qy      244 CTCGACCATCACTATTACAAGAACCCTTTTCCAAAAGAAGAAAGGGATCTCTACTCCACA 303
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      138 LeuAspHisHisTyrTyrLysAsnProPheProLysGluGluArgAspLeuTyrSerThr 157

Qy      304 TCTCTTGCATTTAGGCTCCTCAGAGAACATGGTTTTCAAGTCGCACAAGAGGTATTCGAT 363
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      158 SerLeuAlaPheArgLeuLeuArgGluHisGlyPheGlnValAlaGlnGluValPheAsp 177

Qy      364 AGTTTCAAGAACGAGGAGGGTGAGTTCAAAGAAAGCCTTAGCGACGACACCAGAGGATTG 423
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      178 SerPheLysAsnGluGluGlyGluPheLysGluSerLeuSerAspAspThrArgGlyLeu 197

Qy      424 TTGCAACTGTATGAAGCTTCCTTTCTGTTGACGGAAGGCGAAACCACGCTCGAGTCAGCG 483
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      198 LeuGlnLeuTyrGluAlaSerPheLeuLeuThrGluGlyGluThrThrLeuGluSerAla 217

Qy      484 AGGGAATTCGCCACCAAATTTTTGGAGGAAAAAGTGAACGAGGGTGGTGTTGATGGCGAC 543
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      218 ArgGluPheAlaThrLysPheLeuGluGluLysValAsnGluGlyGlyValAspGlyAsp 237

Qy      544 CTTTTAACAAGAATCGCATATTCTTTGGACATCCCTCTTCATTGGAGGATTAAAAGGCCA 603
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      238 LeuLeuThrArgIleAlaTyrSerLeuAspIleProLeuHisTrpArgIleLysArgPro 257

Qy      604 AATGCACCTGTGTGGATCGAATGGTATAGGAAGAGGCCCGACATGAATCCAGTAGTGTTG 663
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      258 AsnAlaProValTrpIleGluTrpTyrArgLysArgProAspMetAsnProValValLeu 277

Qy      664 GAGCTTGCCATACTCGACTTAAATATTGTTCAAGCACAATTTCAAGAAGAGCTCAAAGAA 723
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      278 GluLeuAlaIleLeuAspLeuAsnIleValGlnAlaGlnPheGlnGluGluLeuLysGlu 297

Qy      724 TCCTTCAGGTGGTGGAGAAATACTGGGTTTGTTGAGAAGCTGCCCTTCGCAAGGGATAGA 783
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      298 SerPheArgTrpTrpArgAsnThrGlyPheValGluLysLeuProPheAlaArgAspArg 317

Qy      784 CTGGTGGAATGCTACTTTTGGAATACTGGGATCATCGAGCCACGTCAGCATGCAAGTGCA 843
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      318 LeuValGluCysTyrPheTrpAsnThrGlyIleIleGluProArgGlnHisAlaSerAla 337

Qy      844 AGGATAATGATGGGCAAAGTCAACGCTCTGATTACGGTGATCGATGATATTTATGATGTC 903
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      338 ArgIleMetMetGlyLysValAsnAlaLeuIleThrValIleAspAspIleTyrAspVal 357

Qy      904 TATGGCACCTTAGAAGAACTCGAACAATTCACTGACCTCATTCGAAGATGGGATATAAAC 963
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      358 TyrGlyThrLeuGluGluLeuGluGlnPheThrAspLeuIleArgArgTrpAspIleAsn 377

Qy      964 TCAATCGACCAACTTCCCGATTACATGCAACTGTGCTTTCTTGCACTCAACAACTTCGTC 1023
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      378 SerIleAspGlnLeuProAspTyrMetGlnLeuCysPheLeuAlaLeuAsnAsnPheVal 397

Qy     1024 GATGATACATCGTACGATGTTATGAAGGAGAAAGGCGTCAACGTTATACCCTACCTGCGG 1083
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      398 AspAspThrSerTyrAspValMetLysGluLysGlyValAsnValIleProTyrLeuArg 417

Qy     1084 CAATCGTGGGTTGATTTGGCGGATAAGTATATGGTAGAGGCACGGTGGTTCTACGGCGGG 1143
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      418 GlnSerTrpValAspLeuAlaAspLysTyrMetValGluAlaArgTrpPheTyrGlyGly 437

Qy     1144 CACAAACCAAGTTTGGAAGAGTATTTGGAGAACTCATGGCAGTCGATAAGTGGGCCCTGT 1203
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      438 HisLysProSerLeuGluGluTyrLeuGluAsnSerTrpGlnSerIleSerGlyProCys 457

Qy     1204 ATGTTAACGCACATATTCTTCCGAGTAACAGATTCGTTCACAAAGGAGACCGTCGACAGT 1263
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      458 MetLeuThrHisIlePhePheArgValThrAspSerPheThrLysGluThrValAspSer 477

Qy     1264 TTGTACAAATACCACGATTTAGTTCGTTGGTCATCCTTCGTTCTGCGGCTTGCTGATGAT 1323
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      478 LeuTyrLysTyrHisAspLeuValArgTrpSerSerPheValLeuArgLeuAlaAspAsp 497

Qy     1324 TTGGGAACCTCGGTGGAAGAGGTGAGCAGAGGGGATGTGCCGAAATCACTTCAGTGCTAC 1383
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      498 LeuGlyThrSerValGluGluValSerArgGlyAspValProLysSerLeuGlnCysTyr 517

Qy     1384 ATGAGTGACTACAATGCATCGGAGGCGGAGGCGCGGAAGCACGTGAAATGGCTGATAGCG 1443
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      518 MetSerAspTyrAsnAlaSerGluAlaGluAlaArgLysHisValLysTrpLeuIleAla 537

Qy     1444 GAGGTGTGGAAGAAGATGAATGCGGAGAGGGTGTCGAAGGATTCTCCATTCGGCAAAGAT 1503
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      538 GluValTrpLysLysMetAsnAlaGluArgValSerLysAspSerProPheGlyLysAsp 557

Qy     1504 TTTATAGGATGTGCAGTTGATTTAGGAAGGATGGCGCAGTTGATGTACCATAATGGAGAT 1563
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      558 PheIleGlyCysAlaValAspLeuGlyArgMetAlaGlnLeuMetTyrHisAsnGlyAsp 577

Qy     1564 GGGCACGGCACACAACACCCTATTATACATCAACAAATGACCAGAACCTTATTCGAGCCC 1623
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      578 GlyHisGlyThrGlnHisProIleIleHisGlnGlnMetThrArgThrLeuPheGluPro 597
```

```
Qy       1624 TTTGCA 1629
              ||||||
Db        598 PheAla 599

RESULT 7
US-09-900-797-22
; Sequence 22, Application US/09900797
; Patent No. 6645762
; GENERAL INFORMATION:
;  APPLICANT: Chappell, Joseph
;  APPLICANT:  No. 66457621, Joseph P.
;  APPLICANT:  Starks, Courtney M.
;  APPLICANT:  Manna, Kathleen R.
;  TITLE OF INVENTION: SYNTHASES
;  FILE REFERENCE: 07678-025001
;  CURRENT APPLICATION NUMBER: US/09/900,797
;  CURRENT FILING DATE:  2001-07-06
;  PRIOR APPLICATION NUMBER: US/09/398,395
;  PRIOR FILING DATE: 1999-09-17
;  PRIOR APPLICATION NUMBER: 60/130,628
;  PRIOR FILING DATE: 1999-04-22
;  PRIOR APPLICATION NUMBER: 60/150,262
;  PRIOR FILING DATE: 1999-08-23
;  NUMBER OF SEQ ID NOS: 58
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 22
;   LENGTH: 599
;   TYPE: PRT
;   ORGANISM: Mentha spicata
US-09-900-797-22

Alignment Scores:
Pred. No.:              2.52e-284     Length:         599
Score:                  2876.00       Matches:        542
Percent Similarity:     100.00%       Conservative:   0
Best Local Similarity:  100.00%       Mismatches:     0
Query Match:            98.90%        Indels:         0
DB:                     4             Gaps:           0

US-09-938-956-6 (1-1632) x US-09-900-797-22 (1-599)

Qy          4 AGACGATCCGGAAACTACAACCCTTCTCGTTGGATGTCAACTTCATCCAATCGCTTCTC 63
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         58 ArgArgSerGlyAsnTyrAsnProSerArgTrpAspValAsnPheIleGlnSerLeuLeu 77

Qy         64 AGTGACTATAAGGAGGACAAACACGTGATTAGGGCTTCTGAGCTGGTCACTTTGGTGAAG 123
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         78 SerAspTyrLysGluAspLysHisValIleArgAlaSerGluLeuValThrLeuValLys 97

Qy        124 ATGGAACTGGAGAAAGAAACGGATCAAATTCGACAACTTGAGTTGATCGATGACTTGCAG 183
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         98 MetGluLeuGluLysGluThrAspGlnIleArgGlnLeuGluLeuIleAspAspLeuGln 117

Qy        184 AGGATGGGGCTGTCCGATCATTTCCAAAATGAGTTCAAAGAAATCTTGTCCTCTATATAT 243
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        118 ArgMetGlyLeuSerAspHisPheGlnAsnGluPheLysGluIleLeuSerSerIleTyr 137

Qy        244 CTCGACCATCACTATTACAAGAACCCTTTTCCAAAAGAAGAAAGGGATCTCTACTCCACA 303
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        138 LeuAspHisHisTyrTyrLysAsnProPheProLysGluGluArgAspLeuTyrSerThr 157

Qy        304 TCTCTTGCATTTAGGCTCCTCAGAGAACATGGTTTTCAAGTCGCACAAGAGGTATTCGAT 363
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        158 SerLeuAlaPheArgLeuLeuArgGluHisGlyPheGlnValAlaGlnGluValPheAsp 177

Qy        364 AGTTTCAAGAACGAGGAGGGTGAGTTCAAAGAAAGCCTTAGCGACGACACCAGAGGATTG 423
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        178 SerPheLysAsnGluGluGlyGluPheLysGluSerLeuSerAspAspThrArgGlyLeu 197

Qy        424 TTGCAACTGTATGAAGCTTCCTTTCTGTTGACGGAAGGCGAAACCACGCTCGAGTCAGCG 483
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        198 LeuGlnLeuTyrGluAlaSerPheLeuLeuThrGluGlyGluThrThrLeuGluSerAla 217

Qy        484 AGGGAATTCGCCACCAAATTTTTGGAGGAAAAAGTGAACGAGGGTGGTGTTGATGGCGAC 543
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        218 ArgGluPheAlaThrLysPheLeuGluGluLysValAsnGluGlyGlyValAspGlyAsp 237

Qy        544 CTTTTAACAAGAATCGCATATTCTTTGGACATCCCTCTTCATTGGAGGATTAAAAGGCCA 603
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        238 LeuLeuThrArgIleAlaTyrSerLeuAspIleProLeuHisTrpArgIleLysArgPro 257

Qy        604 AATGCACCTGTGTGGATCGAATGGTATAGGAAGAGGCCCGACATGAATCCAGTAGTGTTG 663
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        258 AsnAlaProValTrpIleGluTrpTyrArgLysArgProAspMetAsnProValValLeu 277

Qy        664 GAGCTTGCCATACTCGACTTAAATATTGTTCAAGCACAATTTCAAGAAGAGCTCAAAGAA 723
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        278 GluLeuAlaIleLeuAspLeuAsnIleValGlnAlaGlnPheGlnGluGluLeuLysGlu 297

Qy        724 TCCTTCAGGTGGTGGAGAAATACTGGGTTTGTTGAGAAGCTGCCCTTCGCAAGGGATAGA 783
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        298 SerPheArgTrpTrpArgAsnThrGlyPheValGluLysLeuProPheAlaArgAspArg 317

Qy        784 CTGGTGGAATGCTACTTTTGGAATACTGGGATCATCGAGCCACGTCAGCATGCAAGTGCA 843
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        318 LeuValGluCysTyrPheTrpAsnThrGlyIleIleGluProArgGlnHisAlaSerAla 337

Qy        844 AGGATAATGATGGGCAAAGTCAACGCTCTGATTACGGTGATCGATGATATTTATGATGTC 903
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        338 ArgIleMetMetGlyLysValAsnAlaLeuIleThrValIleAspAspIleTyrAspVal 357

Qy        904 TATGGCACCTTAGAAGAACTCGAACAATTCACTGACCTCATTCGAAGATGGGATATAAAC 963
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        358 TyrGlyThrLeuGluGluLeuGluGlnPheThrAspLeuIleArgArgTrpAspIleAsn 377

Qy        964 TCAATCGACCAACTTCCCGATTACATGCAACTGTGCTTTCTTGCACTCAACAACTTCGTC 1023
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        378 SerIleAspGlnLeuProAspTyrMetGlnLeuCysPheLeuAlaLeuAsnAsnPheVal 397

Qy       1024 GATGATACATCGTACGATGTTATGAAGGAGAAAGGCGTCAACGTTATACCCTACCTGCGG 1083
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        398 AspAspThrSerTyrAspValMetLysGluLysGlyValAsnValIleProTyrLeuArg 417

Qy       1084 CAATCGTGGGTTGATTTGGCGGATAAGTATATGGTAGAGGCACGGTGGTTCTACGGCGGG 1143
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        418 GlnSerTrpValAspLeuAlaAspLysTyrMetValGluAlaArgTrpPheTyrGlyGly 437

Qy       1144 CACAAACCAAGTTTGGAAGAGTATTTGGAGAACTCATGGCAGTCGATAAGTGGGCCCTGT 1203
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        438 HisLysProSerLeuGluGluTyrLeuGluAsnSerTrpGlnSerIleSerGlyProCys 457

Qy       1204 ATGTTAACGCACATATTCTTCCGAGTAACAGATTCGTTCACAAAGGAGACCGTCGACAGT 1263
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        458 MetLeuThrHisIlePhePheArgValThrAspSerPheThrLysGluThrValAspSer 477

Qy       1264 TTGTACAAATACCACGATTTAGTTCGTTGGTCATCCTTCGTTCTGCGGCTTGCTGATGAT 1323
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        478 LeuTyrLysTyrHisAspLeuValArgTrpSerSerPheValLeuArgLeuAlaAspAsp 497

Qy       1324 TTGGGAACCTCGGTGGAAGAGGTGAGCAGAGGGGATGTGCCGAAATCACTTCAGTGCTAC 1383
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        498 LeuGlyThrSerValGluGluValSerArgGlyAspValProLysSerLeuGlnCysTyr 517

Qy       1384 ATGAGTGACTACAATGCATCGGAGGCGGAGGCGCGGAAGCACGTGAAATGGCTGATAGCG 1443
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        518 MetSerAspTyrAsnAlaSerGluAlaGluAlaArgLysHisValLysTrpLeuIleAla 537

Qy       1444 GAGGTGTGGAAGAAGATGAATGCGGAGAGGGTGTCGAAGGATTCTCCATTCGGCAAAGAT 1503
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        538 GluValTrpLysLysMetAsnAlaGluArgValSerLysAspSerProPheGlyLysAsp 557

Qy       1504 TTTATAGGATGTGCAGTTGATTTAGGAAGGATGGCGCAGTTGATGTACCATAATGGAGAT 1563
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        558 PheIleGlyCysAlaValAspLeuGlyArgMetAlaGlnLeuMetTyrHisAsnGlyAsp 577

Qy       1564 GGGCACGGCACACAACACCCTATTATACATCAACAAATGACCAGAACCTTATTCGAGCCC 1623
```

```
Db        578 GlyHisGlyThrGlnHisProIleIleHisGlnGlnMetThrArgThrLeuPheGluPro 597

Qy       1624 TTTGCA 1629

Db        598 PheAla 599

RESULT 8
US-08-937-540-4
; Sequence 4, Application US/08937540
; Patent No. 5891697
;  GENERAL INFORMATION:
;    APPLICANT:  Croteau, Rodney B
;    APPLICANT:  Wise, Mitchell L
;    APPLICANT:  Savage, Thomas J
;    APPLICANT:  Katahira, Eva J
;    TITLE OF INVENTION:  Monoterpene Synthases from Common Sage
;    TITLE OF INVENTION:  (Salvia officinalis)
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  CHRISTENSEN, O'CONNOR, JOHNSON & KINDNESS
;      STREET:  1420 FIFTH AVENUE
;      CITY:  SEATTLE
;      STATE:  WASHINGTON
;      COUNTRY:  USA
;      ZIP:  98101-2347
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/937,540
;      FILING DATE:
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Shelton, Dennis K
;      REGISTRATION NUMBER:  26,997
;      REFERENCE/DOCKET NUMBER:  WSUR111254
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  206 695 1718
;      TELEFAX:  206 224 0779
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  591 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-937-540-4

Alignment Scores:
Pred. No.:              1.26e-157      Length:         591
Score:                  1633.50        Matches:        305
Percent Similarity:     75.96%         Conservative:   109
Best Local Similarity:  55.96%         Mismatches:     116
Query Match:            56.17%         Indels:         15
DB:                     2              Gaps:           7

US-09-938-956-6 (1-1632) x US-08-937-540-4 (1-591)

Qy          4 AGACGATCCGGAAACTACAACCCTTCTCGTTGGGATGTCAACTTCATCCAATCGCTTCTC 63

Db         57 ArgArgThrGlyGlyTyrGlnProThrLeuTrpAspPheSerThrIleGlnLeuPheAsp 76

Qy         64 AGTGACTATAAGGAGGACAAACACGTGATTAGGGCTTCTGAGCTGGTCACTTTGGTGAAG 123

Db         77 SerGluTyrLysGluGluLysHisLeuMetArgAlaAlaGlyMetIleAlaGlnValAsn 96

Qy        124 ATGGAACTGGAGAAAGAAACGGATCAAATTCGACAACTTGAGTTGATCGATGACTTGCAG 183

Db         97 MetLeuLeuGlnGluGluValAspSerIleGlnArgLeuGluLeuIleAspAspLeuArg 116

Qy        184 AGGATGGGGCTGTCCGATCATTTCCAAAATGAGTTCAAAGAAATCTTGTCCTCTATATAT 243

Db        117 ArgLeuGlyIleSerCysHisPheAspArgGluIleValGluIleLeuAsnSerLysTyr 136

Qy        244 CTCGACCATCACTATTACAAGAACCCTTTTCCAAAAGAAGAAAGGGATCTCTACTCCACA 303

Db        137 ------------TyrThrAsnAsnGluIle------AspGluSerAspLeuTyrSerThr 150

Qy        304 TCTCTTGCATTTAGGCTCCTCAGAGAACATGGTTTTCAAGTCGCACAAGAGGTATTCGAT 363

Db        151 AlaLeuArgPheLysLeuLeuArgGlnTyrAspPheSerValSerGlnGluValPheAsp 170

Qy        364 AGTTTCAAGAACGAGGAGGGT---GAGTTCAAAGAAAGCCTTAGCGACGACACCAGAGGA 420

Db        171 CysPheLysAsnAspLysGlyThrAspPheLysProSerLeuValAspAspThrArgGly 190

Qy        421 TTGTTGCAACTGTATGAAGCTTCCTTTCTGTTGACGGAAGGCGAAACCACGCTCGAGTCA 480

Db        191 LeuLeuGlnLeuTyrGluAlaSerPheLeuSerAlaGlnGlyGluGluThrLeuHisLeu 210

Qy        481 GCGAGGGAATTCGCCACCAAATTTTTGGAGGAAAAAGTGAACGAGGGTGGTGTTGATGGC 540

Db        211 AlaArgAspPheAlaThrLysPheLeuHisLysArgValLeu---------ValAspLys 227

Qy        541 GAC------CTTTTAACAAGAATCGCATATTCTTTGGACATCCCTCTTCATTGGAGGATT 594

Db        228 AspIleAsnLeuLeuSerSerIleGluArgAlaLeuGluLeuProThrHisTrpArgVal 247

Qy        595 AAAAGGCCAAATGCACCTGTGTGGATCGAATGGTATAGGAAGAGGCCCGACATGAATCCA 654

Db        248 GlnMetProAsnAlaArgSerPheIleAspAlaTyrLysArgArgProAspMetAsnPro 267

Qy        655 GTAGTGTTGGAGCTTGCCATACTCGACTTAAATATTGTTCAAGCACAATTTCAAGAAGAG 714

Db        268 ThrValLeuGluLeuAlaLysLeuAspPheAsnMetValGlnAlaGlnPheGlnGlnGlu 287

Qy        715 CTCAAAGAATCCTTCAGGTGGTGGAGAAATACTGGGTTTGTTGAGAAGCTGCCCTTCGCA 774

Db        288 LeuLysGluAlaSerArgTrpTrpAsnSerThrGlyLeuValHisGluLeuProPheVal 307

Qy        775 AGGGATAGACTGGTGGAATGCTACTTTTGGAATACTGGGATCATCGAGCCACGTCAGCAT 834

Db        308 ArgAspArgIleValGluCysTyrTyrTrpThrThrGlyValValGluArgArgGluHis 327

Qy        835 GCAAGTGCAAGGATAATGATGGGCAAAGTCAACGCTCTGATTACGGTGATCGATGATATT 894

Db        328 GlyTyrGluArgIleMetLeuThrLysIleAsnAlaLeuValThrThrIleAspAspVal 347

Qy        895 TATGATGTCTATGGCACCTTAGAAGAACTCGAACAATTCACTGACCTCATTCGAAGATGG 954

Db        348 PheAspIleTyrGlyThrLeuGluGluLeuGlnLeuPheThrThrAlaIleGlnArgTrp 367

Qy        955 GATATAAACTCAATCGACCAACTTCCCGATTACATGCAACTGTGCTTTCTTGCACTCAAC 1014

Db        368 AspIleGluSerMetLysGlnLeuProProTyrMetGlnIleCysTyrLeuAlaLeuPhe 387

Qy       1015 AACTTCGTCGATGATACATCGTACGATGTTATGAAGGAGAAAGGCGTCAACGTTATACCC 1074

Db        388 AsnPheValAsnGluMetAlaTyrAspThrLeuArgAspLysGlyPheAsnSerThrPro 407

Qy       1075 TACCTGCGGCAATCGTGGGTTGATTTGGCGGATAAGTATATGGTAGAGGCACGGTGGTTC 1134

Db        408 TyrLeuArgLysAlaTrpValAspLeuValGluSerTyrLeuIleGluAlaLysTrpTyr 427

Qy       1135 TACGGCGGGCACAAACCAAGTTTGGAAGAGTATTTGGAGAACTCATGGCAGTCGATAAGT 1194

Db        428 TyrMetGlyHisLysProSerLeuGluGluTyrMetLysAsnSerTrpIleSerIleGly 447

Qy       1195 GGGCCCTGTATGTTAACGCACATATTCTTCCGAGTAACAGATTCGTTCACAAAGGAGACC 1254

Db        448 GlyIleProIleLeuSerHisLeuPhePheArgLeuThrAspSerIleGluGluGluAsp 467

Qy       1255 GTCGACAGTTTGTACAAATACCACGATTTAGTTCGTTGGTCATCCTTCGTTCTGCGGCTT 1314
```

```
Db        468 AlaGluSerMetHisLysTyrHisAspIleValArgAlaSerCysThrIleLeuArgLeu 487

Qy       1315 GCTGATGATTTGGGAACCTCGGTGGAAGAGGTGAGCAGAGGGGATGTGCCGAAATCACTT 1374
              ||||||||:::|||||||||::::::||||||   |||||||||||||||||||:::
Db        488 AlaAspAspMetGlyThrSerLeuAspGluValGluArgGlyAspValProLysSerVal 507

Qy       1375 CAGTGCTACATGAGTGACTACAATGCATCGGAGGCGGAGGCGCGGAAGCACGTGAAATGG 1434
              |||||||||||::::::   |||||||||||   ||||||||:::||||||:::
Db        508 GlnCysTyrMetAsnGluLysAsnAlaSerGluGluGluAlaArgGluHisValArgSer 527

Qy       1435 CTGATAGCGGAGGTGTGGAAGAAGATGAATGCGGAGAGGGTGTCGAAGGATTCTCCATTC 1494
              ||||||   :::  ||||||  ||||||  |||   ::::::         |||  |||
Db        528 LeuIleAspGlnThrTrpLysMetMetAsnLysGluMetMetThr------SerSerPhe 545

Qy       1495 GGCAAAGATTTTATAGGATGTGCAGTTGATTTAGGAAGGATGGCGCAGTTGATGTAC--- 1551
                 |||  |||:::     :::  :::|||  |||||||||||   :::|||
Db        546 SerLysTyrPheValGlnValSerAlaAsnLeuAlaArgMetAlaGlnTrpIleTyrGln 565

Qy       1552 CATAATGGAGATGGGCACGGCACACAACACCCTATTATACATCAACAAATGACCAGAACC 1611
              |||      ||||||  |||  ||||||   :::::::::::   :::
Db        566 HisGluSerAspGlyPheGlyMetGlnHisSerLeuValAsnLysMetLeuArgGlyLeu 585

Qy       1612 TTATTCGAGCCCTTT 1626
              ||||||:::   :::
Db        586 LeuPheAspArgTyr 590

RESULT 9
US-09-398-395A-24
; Sequence 24, Application US/09398395A
; Patent No. 6468772
; GENERAL INFORMATION:
;  APPLICANT: Chappell, Joseph
;  APPLICANT:  No. 64687721, Joseph P.
;  APPLICANT:  Starks, Courtney M.
;  APPLICANT:  Manna, Kathleen R.
;  TITLE OF INVENTION: SYNTHASES
;  FILE REFERENCE: 07678-025001
;  CURRENT APPLICATION NUMBER: US/09/398,395A
;  CURRENT FILING DATE:  1999-09-17
;  PRIOR APPLICATION NUMBER: 60/100,993
;  PRIOR FILING DATE: 1998-09-18
;  PRIOR APPLICATION NUMBER: 60/130,628
;  PRIOR FILING DATE: 1999-04-22
;  PRIOR APPLICATION NUMBER: 60/150,262
;  PRIOR FILING DATE: 1999-08-23
;  NUMBER OF SEQ ID NOS: 58
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 24
;   LENGTH: 591
;   TYPE: PRT
;   ORGANISM: Salvia officinalis
US-09-398-395A-24

Alignment Scores:
Pred. No.:             1.26e-157     Length:        591
Score:                 1633.50       Matches:       305
Percent Similarity:    75.96%        Conservative:  109
Best Local Similarity: 55.96%        Mismatches:    116
Query Match:           56.17%        Indels:        15
DB:                    4             Gaps:          7

US-09-938-956-6 (1-1632) x US-09-398-395A-24 (1-591)

Qy          4 AGACGATCCGGAAACTACAACCCTTCTCGTTGGGATGTCAACTTCATCCAATCGCTTCTC 63
              ||||||:::|||   |||   |||:::   ||||||   :::   ||||||
Db         57 ArgArgThrGlyGlyTyrGlnProThrLeuTrpAspPheSerThrIleGlnLeuPheAsp 76

Qy         64 AGTGACTATAAGGAGGACAAACACGTGATTAGGGCTTCTGAGCTGGTCACTTTGGTGAAG 123
              |||:::|||||||||:::||||||::::::||||||:::   ::::::       |||
Db         77 SerGluTyrLysGluGluLysHisLeuMetArgAlaAlaGlyMetIleAlaGlnValAsn 96

Qy        124 ATGGAACTGGAGAAAGAAACGGATCAAATTCGACAACTTGAGTTGATCGATGACTTGCAG 183
              |||   |||::::::|||   |||   |||::::::|||||||||||||||||||||:::
Db         97 MetLeuLeuGlnGluGluValAspSerIleGlnArgLeuGluLeuIleAspAspLeuArg 116

Qy        184 AGGATGGGGCTGTCCGATCATTTCCAAAATGAGTTCAAAGAAATCTTGTCCTCTATATAT 243
              |||:::|||:::|||   ||||||      |||      |||||||||:::|||   |||
Db        117 ArgLeuGlyIleSerCysHisPheAspArgGluIleValGluIleLeuAsnSerLysTyr 136

Qy        244 CTCGACCATCACTATTACAAGAACCCTTTTCCAAAAGAAGAAAGGGATCTCTACTCCACA 303
                          |||      |||            :::|||   |||||||||||||||
Db        137 ------------TyrThrAsnAsnGluIle------AspGluSerAspLeuTyrSerThr 150

Qy        304 TCTCTTGCATTTAGGCTCCTCAGAGAACATGGTTTTCAAGTCGCACAAGAGGTATTCGAT 363
              :::|||   |||:::||||||||||||::::::   |||      |||:::|||||||||||||
Db        151 AlaLeuArgPheLysLeuLeuArgGlnTyrAspPheSerValSerGlnGluValPheAsp 170

Qy        364 AGTTTCAAGAACGAGGAGGGT---GAGTTCAAAGAAAGCCTTAGCGACGACACCAGAGGA 420
                 |||||||||::::::|||   :::||||||   ||||||   ||||||||||||||||
Db        171 CysPheLysAsnAspLysGlyThrAspPheLysProSerLeuValAspAspThrArgGly 190

Qy        421 TTGTTGCAACTGTATGAAGCTTCCTTTCTGTTGACGGAAGGCGAAACCACGCTCGAGTCA 480
              |||||||||||||||||||||||||||       :::||||||   ||||||
Db        191 LeuLeuGlnLeuTyrGluAlaSerPheLeuSerAlaGlnGlyGluGluThrLeuHisLeu 210

Qy        481 GCGAGGGAATTCGCCACCAAATTTTTGGAGGAAAAAGTGAACGAGGGTGGTGTTGATGGC 540
              ||||||:::|||||||||||||||||||      ::::::|||            ||||||
Db        211 AlaArgAspPheAlaThrLysPheLeuHisLysArgValLeu---------ValAspLys 227

Qy        541 GAC------CTTTTAACAAGAATCGCATATTCTTTGGACATCCCTCTTCATTGGAGGATT 594
              |||      ||||||:::   |||      :::|||::::::|||   |||||||||:::
Db        228 AspIleAsnLeuLeuSerSerIleGluArgAlaLeuGluLeuProThrHisTrpArgVal 247

Qy        595 AAAAGGCCAAATGCACCTGTGTGGATCGAATGGTATAGGAAGAGGCCCGACATGAATCCA 654
              :::   |||||||||      :::|||:::   |||:::::::||||||||||||||||||
Db        248 GlnMetProAsnAlaArgSerPheIleAspAlaTyrLysArgArgProAspMetAsnPro 267

Qy        655 GTAGTGTTGGAGCTTGCCATACTCGACTTAAATATTGTTCAAGCACAATTTCAAGAAGAG 714
                 |||||||||||||||   ||||||   |||:::||||||||||||||||||:::|||
Db        268 ThrValLeuGluLeuAlaLysLeuAspPheAsnMetValGlnAlaGlnPheGlnGlnGlu 287

Qy        715 CTCAAAGAATCCTTCAGGTGGTGGAGAAATACTGGGTTTGTTGAGAAGCTGCCCTTCGCA 774
              |||||||||:::   |||||||||   :::||||||   |||   :::|||||||||
Db        288 LeuLysGluAlaSerArgTrpTrpAsnSerThrGlyLeuValHisGluLeuProPheVal 307

Qy        775 AGGGATAGACTGGTGGAATGCTACTTTTGGAATACTGGGATCATCGAGCCACGTCAGCAT 834
              |||||||||:::|||||||||||||:::|||   ||||||::::::|||   |||:::|||
Db        308 ArgAspArgIleValGluCysTyrTyrTrpThrThrGlyValValGluArgArgGluHis 327

Qy        835 GCAAGTGCAAGGATAATGATGGGCAAAGTCAACGCTCTGATTACGGTGATCGATGATATT 894
                       |||||||||:::   |||:::||||||||||:::|||   |||||||||:::
Db        328 GlyTyrGluArgIleMetLeuThrLysIleAsnAlaLeuValThrThrIleAspAspVal 347

Qy        895 TATGATGTCTATGGCACCTTAGAAGAACTCGAACAATTCACTGACCTCATTCGAAGATGG 954
              :::|||:::||||||||||||||||||||:::   ||||||      |||:::||||||
Db        348 PheAspIleTyrGlyThrLeuGluGluLeuGlnLeuPheThrThrAlaIleGlnArgTrp 367

Qy        955 GATATAAACTCAATCGACCAACTTCCCGATTACATGCAACTGTGCTTTCTTGCACTCAAC 1014
              ||||||   |||:::   |||||||||   |||||||||:::|||:::|||||||||
Db        368 AspIleGluSerMetLysGlnLeuProProTyrMetGlnIleCysTyrLeuAlaLeuPhe 387

Qy       1015 AACTTCGTCGATGATACATCGTACGATGTTATGAAGGAGAAAGGCGTCAACGTTATACCC 1074
              |||||||||::::::   :::||||||   :::::::::||||||   |||      |||
Db        388 AsnPheValAsnGluMetAlaTyrAspThrLeuArgAspLysGlyPheAsnSerThrPro 407

Qy       1075 TACCTGCGGCAATCGTGGGTTGATTTGGCGGATAAGTATATGGTAGAGGCACGGTGGTTC 1134
              |||||||||::::::|||||||||||||   :::   |||::::::||||||:::|||:::
Db        408 TyrLeuArgLysAlaTrpValAspLeuValGluSerTyrLeuIleGluAlaLysTrpTyr 427

Qy       1135 TACGGCGGGCACAAACCAAGTTTGGAAGAGTATTTGGAGAACTCATGGCAGTCGATAAGT 1194
              |||   |||||||||||||||||||||||||||::::::|||||||||   ||||||
Db        428 TyrMetGlyHisLysProSerLeuGluGluTyrMetLysAsnSerTrpIleSerIleGly 447

Qy       1195 GGGCCCTGTATGTTAACGCACATATTCTTCCGAGTAACAGATTCGTTCACAAAGGAGACC 1254
              |||      :::|||:::|||:::|||||||||:::|||||||||      :::|||
Db        448 GlyIleProIleLeuSerHisLeuPhePheArgLeuThrAspSerIleGluGluAsp 467
```

```
Qy       1255 GTCGACAGTTTGTACAAATACCACGATTTAGTTCGTTGGTCATCCTTCGTTCTGCGGCTT 1314
Db        468 AlaGluSerMetHisLysTyrHisAspIleValArgAlaSerCysThrIleLeuArgLeu 487

Qy       1315 GCTGATGATTTGGGAACCTCGGTGGAAGAGGTGAGCAGAGGGGATGTGCCGAAATCACTT 1374
Db        488 AlaAspAspMetGlyThrSerLeuAspGluValGluArgGlyAspValProLysSerVal 507

Qy       1375 CAGTGCTACATGAGTGACTACAATGCATCGGAGGCGGAGGCGCGGAAGCACGTGAAATGG 1434
Db        508 GlnCysTyrMetAsnGluLysAsnAlaSerGluGluGluAlaArgGluHisValArgSer 527

Qy       1435 CTGATAGCGGAGGTGTGGAAGAAGATGAATGCGGAGAGGGTGTCGAAGGATTCTCCATTC 1494
Db        528 LeuIleAspGlnThrTrpLysMetMetAsnLysGluMetMetThr------SerSerPhe 545

Qy       1495 GGCAAAGATTTTATAGGATGTGCAGTTGATTTAGGAAGGATGGCGCAGTTGATGTAC--- 1551
Db        546 SerLysTyrPheValGlnValSerAlaAsnLeuAlaArgMetAlaGlnTrpIleTyrGln 565

Qy       1552 CATAATGGAGATGGGCACGGCACACAACACCCTATTATACATCAACAAATGACCAGAACC 1611
Db        566 HisGluSerAspGlyPheGlyMetGlnHisSerLeuValAsnLysMetLeuArgGlyLeu 585

Qy       1612 TTATTCGAGCCCTTT 1626
Db        586 LeuPheAspArgTyr 590
```

RESULT 10
US-09-887-586A-24
; Sequence 24, Application US/09887586A
; Patent No. 6495354
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. 64953541, Joseph P.
; APPLICANT: Starks, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/887,586A
; CURRENT FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: 09/398,395
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 24
; LENGTH: 591
; TYPE: PRT
; ORGANISM: Salvia officinalis
US-09-887-586A-24

Alignment Scores:

Pred. No.:	1.26e-157	Length:	591
Score:	1633.50	Matches:	305
Percent Similarity:	75.96%	Conservative:	109
Best Local Similarity:	55.96%	Mismatches:	116
Query Match:	56.17%	Indels:	15
DB:	4	Gaps:	7

US-09-938-956-6 (1-1632) x US-09-887-586A-24 (1-591)

```
Qy          4 AGACGATCCGGAAACTACAACCCTTCTCGTTGGGATGTCAACTTCATCCAATCGCTTCTC 63
Db         57 ArgArgThrGlyGlyTyrGlnProThrLeuTrpAspPheSerThrIleGlnLeuPheAsp 76

Qy         64 AGTGACTATAAGGAGGACAAACACGTGATTAGGGCTTCTGAGCTGGTCACTTTGGTGAAG 123
Db         77 SerGluTyrLysGluGluLysHisLeuMetArgAlaAlaGlyMetIleAlaGlnValAsn 96

Qy        124 ATGGAACTGGAGAAAGAAACGGATCAAATTCGACAACTTGAGTTGATCGATGACTTGCAG 183
Db         97 MetLeuLeuGlnGluGluValAspSerIleGlnArgLeuGluLeuIleAspAspLeuArg 116

Qy        184 AGGATGGGGCTGTCCGATCATTTCCAAAATGAGTTCAAAGAAATCTTGTCCTCTATATAT 243
Db        117 ArgLeuGlyIleSerCysHisPheAspArgGluIleValGluIleLeuAsnSerLysTyr 136

Qy        244 CTCGACCATCACTATTACAAGAACCCTTTTCCAAAAGAAGAAAGGGATCTCTACTCCACA 303
Db        137 ------------TyrThrAsnAsnGluIle------AspGluSerAspLeuTyrSerThr 150

Qy        304 TCTCTTGCATTTAGGCTCCTCAGAGAACATGGTTTTCAAGTCGCACAAGAGGTATTCGAT 363
Db        151 AlaLeuArgPheLysLeuLeuArgGlnTyrAspPheSerValSerGlnGluValPheAsp 170

Qy        364 AGTTTCAAGAACGAGGAGGGT---GAGTTCAAAGAAAGCCTTAGCGACGACACCAGAGGA 420
Db        171 CysPheLysAsnAspLysGlyThrAspPheLysProSerLeuValAspAspThrArgGly 190

Qy        421 TTGTTGCAACTGTATGAAGCTTCCTTTCTGTTGACGGAAGGCGAAACCACGCTCGAGTCA 480
Db        191 LeuLeuGlnLeuTyrGluAlaSerPheLeuSerAlaGlnGlyGluGluThrLeuHisLeu 210

Qy        481 GCGAGGGAATTCGCCACCAAATTTTTGGAGGAAAAAGTGAACGAGGGTGGTGTTGATGGC 540
Db        211 AlaArgAspPheAlaThrLysPheLeuHisLysArgValLeu---------ValAspLys 227

Qy        541 GAC------CTTTTAACAAGAATCGCATATTCTTTGGACATCCCTCTTCATTGGAGGATT 594
Db        228 AspIleAsnLeuLeuSerSerIleGluArgAlaLeuGluLeuProThrHisTrpArgVal 247

Qy        595 AAAAGGCCAAATGCACCTGTGTGGATCGAATGGTATAGGAAGAGGCCCGACATGAATCCA 654
Db        248 GlnMetProAsnAlaArgSerPheIleAspAlaTyrLysArgArgProAspMetAsnPro 267

Qy        655 GTAGTGTTGGAGCTTGCCATACTCGACTTAAATATTGTTCAAGCACAATTTCAAGAAGAG 714
Db        268 ThrValLeuGluLeuAlaLysLeuAspPheAsnMetValGlnAlaGlnPheGlnGlnGlu 287

Qy        715 CTCAAAGAATCCTTCAGGTGGTGGAGAAATACTGGGTTTGTTGAGAAGCTGCCCTTCGCA 774
Db        288 LeuLysGluAlaSerArgTrpTrpAsnSerThrGlyLeuValHisGluLeuProPheVal 307

Qy        775 AGGGATAGACTGGTGGAATGCTACTTTTGGAATACTGGGATCATCGAGCCACGTCAGCAT 834
Db        308 ArgAspArgIleValGluCysTyrTyrTrpThrThrGlyValValGluArgArgGluHis 327

Qy        835 GCAAGTGCAAGGATAATGATGGGCAAAGTCAACGCTCTGATTACGGTGATCGATGATATT 894
Db        328 GlyTyrGluArgIleMetLeuThrLysIleAsnAlaLeuValThrThrIleAspAspVal 347

Qy        895 TATGATGTCTATGGCACCTTAGAAGAACTCGAACAATTCACTGACCTCATTCGAAGATGG 954
Db        348 PheAspIleTyrGlyThrLeuGluGluLeuGlnLeuPheThrThrAlaIleGlnArgTrp 367

Qy        955 GATATAAACTCAATCGACCAACTTCCCGATTACATGCAACTGTGCTTTCTTGCACTCAAC 1014
Db        368 AspIleGluSerMetLysGlnLeuProProTyrMetGlnIleCysTyrLeuAlaLeuPhe 387

Qy       1015 AACTTCGTCGATGATACATCGTACGATGTTATGAAGGAGAAAGGCGTCAACGTTATACCC 1074
Db        388 AsnPheValAsnGluMetAlaTyrAspThrLeuArgAspLysGlyPheAsnSerThrPro 407

Qy       1075 TACCTGCGGCAATCGTGGGTTGATTTGGCGGATAAGTATATGGTAGAGGCACGGTGGTTC 1134
Db        408 TyrLeuArgLysAlaTrpValAspLeuValGluSerTyrLeuIleGluAlaLysTrpTyr 427

Qy       1135 TACGGCGGGCACAAACCAAGTTTGGAAGAGTATTTGGAGAACTCATGGCAGTCGATAAGT 1194
Db        428 TyrMetGlyHisLysProSerLeuGluGluTyrMetLysAsnSerTrpIleSerIleGly 447
```

```
Qy       1195 GGGCCCTGTATGTTAACGCACATATTCTTCCGAGTAACAGATTCGTTCACAAAGGAGACC 1254
              |||        :::|||:::|||:::|||||||||:::|||||||||        :::|||
Db        448 GlyIleProIleLeuSerHisLeuPhePheArgLeuThrAspSerIleGluGluGluAsp 467

Qy       1255 GTCGACAGTTTGTACAAATACCACGATTTAGTTCGTTGGTCATCCTTCGTTCTGCGGCTT 1314
                 :::|||::::::||||||||||||:::||||||   |||      :::|||||||||
Db        468 AlaGluSerMetHisLysTyrHisAspIleValArgAlaSerCysThrIleLeuArgLeu 487

Qy       1315 GCTGATGATTTGGGAACCTCGGTGGAAGAGGTGAGCAGAGGGGATGTGCCGAAATCACTT 1374
              |||||||||:::|||||||||::::::||||||   ||||||||||||||||||||:::
Db        488 AlaAspAspMetGlyThrSerLeuAspGluValGluArgGlyAspValProLysSerVal 507

Qy       1375 CAGTGCTACATGAGTGACTACAATGCATCGGAGGCGGAGGCGCGGAAGCACGTGAAATGG 1434
              ||||||||||||::::::   ||||||||||||   |||||||||:::||||||:::
Db        508 GlnCysTyrMetAsnGluLysAsnAlaSerGluGluGluAlaArgGluHisValArgSer 527

Qy       1435 CTGATAGCGGAGGTGTGGAAGAAGATGAATGCGGAGAGGGTGTCGAAGGATTCTCCATTC 1494
              ||||||   :::   ||||||   ||||||   |||   ::::::      |||  |||
Db        528 LeuIleAspGlnThrTrpLysMetMetAsnLysGluMetMetThr------SerSerPhe 545

Qy       1495 GGCAAAGATTTTATAGGATGTGCAGTTGATTTAGGAAGGATGGCGCAGTTGATGTAC--- 1551
                 |||   |||:::      :::   :::|||   ||||||||||||   :::|||
Db        546 SerLysTyrPheValGlnValSerAlaAsnLeuAlaArgMetAlaGlnTrpIleTyrGln 565

Qy       1552 CATAATGGAGATGGGCACGGCACACAACACCCTATTATACATCAACAAATGACCAGAACC 1611
              |||      ||||||   |||   ||||||   ::::::::::::   :::
Db        566 HisGluSerAspGlyPheGlyMetGlnHisSerLeuValAsnLysMetLeuArgGlyLeu 585

Qy       1612 TTATTCGAGCCCTTT 1626
              ||||||:::   :::
Db        586 LeuPheAspArgTyr 590
```

RESULT 11

```
US-09-895-752-24
; Sequence 24, Application US/09895752
; Patent No. 6559297
; GENERAL INFORMATION:
;  APPLICANT: Chappell, Joseph
;  APPLICANT:  No. 65592971, Joseph P.
;  APPLICANT:  Starks, Courtney M.
;  APPLICANT:  Manna, Kathleen R.
;  TITLE OF INVENTION: SYNTHASES
;  FILE REFERENCE: 07678-025001
;  CURRENT APPLICATION NUMBER: US/09/895,752
;  CURRENT FILING DATE:  2001-06-29
;  PRIOR APPLICATION NUMBER: 09/398,395
;  PRIOR FILING DATE: 1999-09-17
;  PRIOR APPLICATION NUMBER: 60/100,993
;  PRIOR FILING DATE: 1998-09-18
;  PRIOR APPLICATION NUMBER: 60/130,628
;  PRIOR FILING DATE: 1999-04-22
;  PRIOR APPLICATION NUMBER: 60/150,262
;  PRIOR FILING DATE: 1999-08-23
;  NUMBER OF SEQ ID NOS: 58
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 24
;   LENGTH: 591
;   TYPE: PRT
;   ORGANISM: Salvia officinalis
US-09-895-752-24

Alignment Scores:
Pred. No.:               1.26e-157     Length:        591
Score:                   1633.50       Matches:       305
Percent Similarity:      75.96%        Conservative:  109
Best Local Similarity:   55.96%        Mismatches:    116
Query Match:             56.17%        Indels:        15
DB:                      4             Gaps:          7

US-09-938-956-6 (1-1632) x US-09-895-752-24 (1-591)

Qy          4 AGACGATCCGGAAACTACAACCCTTCTCGTTGGGATGTCAACTTCATCCAATCGCTTCTC 63
               ||||||:::|||   |||   |||:::   ||||||   :::   ||||||
Db         57 ArgArgThrGlyGlyTyrGlnProThrLeuTrpAspPheSerThrIleGlnLeuPheAsp 76

Qy         64 AGTGACTATAAGGAGGACAAACACGTGATTAGGGCTTCTGAGCTGGTCACTTTGGTGAAG 123
              |||:::|||||||||:::||||||::::::||||||:::   ::::::      |||
Db         77 SerGluTyrLysGluGluLysHisLeuMetArgAlaAlaGlyMetIleAlaGlnValAsn 96

Qy        124 ATGGAACTGGAGAAAGAAACGGATCAAATTCGACAACTTGAGTTGATCGATGACTTGCAG 183
              |||   |||::::::|||   |||   |||::::::||||||||||||||||||||:::
Db         97 MetLeuLeuGlnGluGluValAspSerIleGlnArgLeuGluLeuIleAspAspLeuArg 116

Qy        184 AGGATGGGGCTGTCCGATCATTTCCAAAATGAGTTCAAAGAAATCTTGTCCTCTATATAT 243
              |||:::|||:::|||   ||||||   |||      ||||||:::|||   |||
Db        117 ArgLeuGlyIleSerCysHisPheAspArgGluIleValGluIleLeuAsnSerLysTyr 136

Qy        244 CTCGACCATCACTATTACAAGAACCCTTTTCCAAAAGAAGAAAGGGATCTCTACTCCACA 303
                          |||   |||            :::|||   ||||||||||||
Db        137 ------------TyrThrAsnAsnGluIle------AspGluSerAspLeuTyrSerThr 150

Qy        304 TCTCTTGCATTTAGGCTCCTCAGAGAACATGGTTTTCAAGTCGCACAAGAGGTATTCGAT 363
              :::|||   |||:::||||||||||||::::::   |||   |||:::|||||||||||||
Db        151 AlaLeuArgPheLysLeuLeuArgGlnTyrAspPheSerValSerGlnGluValPheAsp 170

Qy        364 AGTTTCAAGAACGAGGAGGGT---GAGTTCAAAGAAAGCCTTAGCGACGACACCAGAGGA 420
              |||||||||::::::|||   :::||||||   ||||||   ||||||||||||||
Db        171 CysPheLysAsnAspLysGlyThrAspPheLysProSerLeuValAspAspThrArgGly 190

Qy        421 TTGTTGCAACTGTATGAAGCTTCCTTTCTGTTGACGGAAGGCGAAACCACGCTCGAGTCA 480
              |||||||||||||||||||||||||||         :::||||||   ||||||
Db        191 LeuLeuGlnLeuTyrGluAlaSerPheLeuSerAlaGlnGlyGluGluThrLeuHisLeu 210

Qy        481 GCGAGGGAATTCGCCACCAAATTTTTGGAGGAAAAAGTGAACGAGGGTGGTGTTGATGGC 540
              ||||||:::||||||||||||||||||   ::::::|||            ||||||
Db        211 AlaArgAspPheAlaThrLysPheLeuHisLysArgValLeu---------ValAspLys 227

Qy        541 GAC------CTTTTAACAAGAATCGCATATTCTTTGGACATCCCTCTTCATTGGAGGATT 594
              |||      ||||||:::   |||   :::|||::::::|||   |||||||||:::
Db        228 AspIleAsnLeuLeuSerSerIleGluArgAlaLeuGluLeuProThrHisTrpArgVal 247

Qy        595 AAAAGGCCAAATGCACCTGTGTGGATCGAATGGTATAGGAAGAGGCCCGACATGAATCCA 654
               :::   |||||||||      :::|||:::   |||::::::||||||||||||||||||
Db        248 GlnMetProAsnAlaArgSerPheIleAspAlaTyrLysArgArgProAspMetAsnPro 267

Qy        655 GTAGTGTTGGAGCTTGCCATACTCGACTTAAATATTGTTCAAGCACAATTTCAAGAAGAG 714
                 |||||||||||||||   ||||||   |||:::|||||||||||||||||||:::|||
Db        268 ThrValLeuGluLeuAlaLysLeuAspPheAsnMetValGlnAlaGlnPheGlnGluGlu 287

Qy        715 CTCAAAGAATCCTTCAGGTGGTGGAGAAATACTGGGTTTGTTGAGAAGCTGCCCTTCGCA 774
              |||||||||:::   |||||||||   |||   :::|||   :::|||||||||
Db        288 LeuLysGluAlaSerArgTrpTrpAsnSerThrGlyLeuValHisGluLeuProPheVal 307

Qy        775 AGGGATAGACTGGTGGAATGCTACTTTTGGAATACTGGGATCATCGAGCCACGTCAGCAT 834
              |||||||||:::||||||||||||:::|||   ||||||::::::|||   |||:::|||
Db        308 ArgAspArgIleValGluCysTyrTyrTrpThrThrGlyValValGluArgArgGluHis 327

Qy        835 GCAAGTGCAAGGATAATGATGGGCAAAGTCAACGCTCTGATTACGGTGATCGATGATATT 894
                       |||||||||:::   |||:::||||||||||||:::|||   |||||||||:::
Db        328 GlyTyrGluArgIleMetLeuThrLysIleAsnAlaLeuValThrThrIleAspAspVal 347

Qy        895 TATGATGTCTATGGCACCTTAGAAGAACTCGAACAATTCACTGACCTCATTCGAAGATGG 954
              :::|||:::||||||||||||||||||||||||:::      ||||||      |||:::||||||
Db        348 PheAspIleTyrGlyThrLeuGluGluLeuGlnLeuPheThrThrAlaIleGlnArgTrp 367

Qy        955 GATATAAACTCAATCGACCAACTTCCCGATTACATGCAACTGTGCTTTCTTGCACTCAAC 1014
              ||||||   |||:::   |||||||||   |||||||||:::|||:::|||||||||
Db        368 AspIleGluSerMetLysGlnLeuProProTyrMetGlnIleCysTyrLeuAlaLeuPhe 387

Qy       1015 AACTTCGTCGATGATACATCGTACGATGTTATGAAGGAGAAAGGCGTCAACGTTATACCC 1074
              |||||||||::::::   :::||||||   :::::::::||||||   |||      |||
Db        388 AsnPheValAsnGluMetAlaTyrAspThrLeuArgAspLysGlyPheAsnSerThrPro 407

Qy       1075 TACCTGCGGCAATCGTGGGTTGATTTGGCGGATAAGTATATGGTAGAGGCACGGTGGTTC 1134
              |||||||||::::::||||||||||||      :::   |||::::::|||||||||:::|||:::
Db        408 TyrLeuArgLysAlaTrpValAspLeuValGluSerTyrLeuIleGluAlaLysTrpTyr 427
```

```
Qy       1135 TACGGCGGGCACAAACCAAGTTTGGAAGAGTATTTGGAGAACTCATGGCAGTCGATAAGT 1194
Db        428 TyrMetGlyHisLysProSerLeuGluGluTyrMetLysAsnSerTrpIleSerIleGly 447

Qy       1195 GGGCCCTGTATGTTAACGCACATATTCTTCCGAGTAACAGATTCGTTCACAAAGGAGACC 1254
Db        448 GlyIleProIleLeuSerHisLeuPhePheArgLeuThrAspSerIleGluGluGluAsp 467

Qy       1255 GTCGACAGTTTGTACAAATACCACGATTTAGTTCGTTGGTCATCCTTCGTTCTGCGGCTT 1314
Db        468 AlaGluSerMetHisLysTyrHisAspIleValArgAlaSerCysThrIleLeuArgLeu 487

Qy       1315 GCTGATGATTTGGGAACCTCGGTGGAAGAGGTGAGCAGAGGGGATGTGCCGAAATCACTT 1374
Db        488 AlaAspAspMetGlyThrSerLeuAspGluValGluArgGlyAspValProLysSerVal 507

Qy       1375 CAGTGCTACATGAGTGACTACAATGCATCGGAGGCGGAGGCGCGGAAGCACGTGAAATGG 1434
Db        508 GlnCysTyrMetAsnGluLysAsnAlaSerGluGluGluAlaArgGluHisValArgSer 527

Qy       1435 CTGATAGCGGAGGTGTGGAAGAAGATGAATGCGGAGAGGGTGTCGAAGGATTCTCCATTC 1494
Db        528 LeuIleAspGlnThrTrpLysMetMetAsnLysGluMetMetThr------SerSerPhe 545

Qy       1495 GGCAAAGATTTTATAGGATGTGCAGTTGATTTAGGAAGGATGGCGCAGTTGATGTAC--- 1551
Db        546 SerLysTyrPheValGlnValSerAlaAsnLeuAlaArgMetAlaGlnTrpIleTyrGln 565

Qy       1552 CATAATGGAGATGGGCACGGCACACAACACCCTATTATACATCAACAAATGACCAGAACC 1611
Db        566 HisGluSerAspGlyPheGlyMetGlnHisSerLeuValAsnLysMetLeuArgGlyLeu 585

Qy       1612 TTATTCGAGCCCTTT 1626
Db        586 LeuPheAspArgTyr 590
```

```
RESULT 12
US-09-903-012B-24
; Sequence 24, Application US/09903012B
; Patent No. 6569656
; GENERAL INFORMATION:
;  APPLICANT: Chappell, Joseph
;  APPLICANT:  No. 65696561, Joseph P.
;  APPLICANT:  Starks, Courtney M.
;  APPLICANT:  Manna, Kathleen R.
;  TITLE OF INVENTION: SYNTHASES
;  FILE REFERENCE: 07678-025001
;  CURRENT APPLICATION NUMBER: US/09/903,012B
;  CURRENT FILING DATE:  2001-07-11
;  PRIOR APPLICATION NUMBER: 60/100,993
;  PRIOR FILING DATE: 1998-09-18
;  PRIOR APPLICATION NUMBER: 60/130,628
;  PRIOR FILING DATE: 1999-04-22
;  PRIOR APPLICATION NUMBER: 60/150,262
;  PRIOR FILING DATE: 1999-08-23
;  NUMBER OF SEQ ID NOS: 58
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 24
;   LENGTH: 591
;   TYPE: PRT
;   ORGANISM: Salvia officinalis
US-09-903-012B-24

Alignment Scores:
Pred. No.:              1.26e-157     Length:         591
Score:                  1633.50       Matches:        305
Percent Similarity:     75.96%        Conservative:   109
Best Local Similarity:  55.96%        Mismatches:     116
Query Match:            56.17%        Indels:         15
DB:                     4             Gaps:           7

US-09-938-956-6 (1-1632) x US-09-903-012B-24 (1-591)
```

```
Qy          4 AGACGATCCGGAAACTACAACCCTTCTCGTTGGGATGTCAACTTCATCCAATCGCTTCTC 63
Db         57 ArgArgThrGlyGlyTyrGlnProThrLeuTrpAspPheSerThrIleGlnLeuPheAsp 76

Qy         64 AGTGACTATAAGGAGGACAAACACGTGATTAGGGCTTCTGAGCTGGTCACTTTGGTGAAG 123
Db         77 SerGluTyrLysGluGluLysHisLeuMetArgAlaAlaGlyMetIleAlaGlnValAsn 96

Qy        124 ATGGAACTGGAGAAAGAAACGGATCAAATTCGACAACTTGAGTTGATCGATGACTTGCAG 183
Db         97 MetLeuLeuGlnGluGluValAspSerIleGlnArgLeuGluLeuIleAspAspLeuArg 116

Qy        184 AGGATGGGGCTGTCCGATCATTTCCAAAATGAGTTCAAAGAAATCTTGTCCTCTATATAT 243
Db        117 ArgLeuGlyIleSerCysHisPheAspArgGluIleValGluIleLeuAsnSerLysTyr 136

Qy        244 CTCGACCATCACTATTACAAGAACCCTTTTCCAAAAGAAGAAAGGATCTCTACTCCACA 303
Db        137 ------------TyrThrAsnAsnGluIle------AspGluSerAspLeuTyrSerThr 150

Qy        304 TCTCTTGCATTTAGGCTCCTCAGAGAACATGGTTTTCAAGTCGCACAAGAGGTATTCGAT 363
Db        151 AlaLeuArgPheLysLeuLeuArgGlnTyrAspPheSerValSerGlnGluValPheAsp 170

Qy        364 AGTTTCAAGAACGAGGAGGGT---GAGTTCAAAGAAAGCCTTAGCGACGACACCAGAGGA 420
Db        171 CysPheLysAsnAspLysGlyThrAspPheLysProSerLeuValAspAspThrArgGly 190

Qy        421 TTGTTGCAACTGTATGAAGCTTCCTTTCTGTTGACGGAAGGCGAAACCACGCTCGAGTCA 480
Db        191 LeuLeuGlnLeuTyrGluAlaSerPheLeuSerAlaGlnGlyGluGluThrLeuHisLeu 210

Qy        481 GCGAGGGAATTCGCCACCAAATTTTTGGAGGAAAAAGTGAACGAGGGTGGTGTTGATGGC 540
Db        211 AlaArgAspPheAlaThrLysPheLeuHisLysArgValLeu---------ValAspLys 227

Qy        541 GAC------CTTTTAACAAGAATCGCATATTCTTTGGACATCCCTCTTCATTGGAGGATT 594
Db        228 AspIleAsnLeuLeuSerSerIleGluArgAlaLeuGluLeuProThrHisTrpArgVal 247

Qy        595 AAAAGGCCAAATGCACCTGTGTGGATCGAATGGTATAGGAAGAGGCCCGACATGAATCCA 654
Db        248 GlnMetProAsnAlaArgSerPheIleAspAlaTyrLysArgArgProAspMetAsnPro 267

Qy        655 GTAGTGTTGGAGCTTGCCATACTCGACTTAAATATTGTTCAAGCACAATTTCAAGAAGAG 714
Db        268 ThrValLeuGluLeuAlaLysLeuAspPheAsnMetValGlnAlaGlnPheGlnGlnGlu 287

Qy        715 CTCAAAGAATCCTTCAGGTGGTGGAGAAATACTGGGTTTGTTGAGAAGCTGCCCTTCGCA 774
Db        288 LeuLysGluAlaSerArgTrpTrpAsnSerThrGlyLeuValHisGluLeuProPheVal 307

Qy        775 AGGGATAGACTGGTGGAATGCTACTTTTGGAATACTGGGATCATCGAGCCACGTCAGCAT 834
Db        308 ArgAspArgIleValGluCysTyrTyrTrpThrThrGlyValValGluArgArgGluHis 327

Qy        835 GCAAGTGCAAGGATAATGATGGGCAAAGTCAACGCTCTGATTACGGTGATCGATGATATT 894
Db        328 GlyTyrGluArgIleMetLeuThrLysIleAsnAlaLeuValThrThrIleAspAspVal 347

Qy        895 TATGATGTCTATGGCACCTTAGAAGAACTCGAACAATTCACTGACCTCATTCGAAGATGG 954
Db        348 PheAspIleTyrGlyThrLeuGluGluLeuGlnLeuPheThrThrAlaIleGlnArgTrp 367

Qy        955 GATATAAACTCAATCGACCAACTTCCCGATTACATGCAACTGTGCTTTCTTGCACTCAAC 1014
Db        368 AspIleGluSerMetLysGlnLeuProProTyrMetGlnIleCysTyrLeuAlaLeuPhe 387

Qy       1015 AACTTCGTCGATGATACATCGTACGATGTTATGAAGGAGAAAGGCGTCAACGTTATACCC 1074
Db        388 AsnPheValAsnGluMetAlaTyrAspThrLeuArgAspLysGlyPheAsnSerThrPro 407
```

```
Qy     1075 TACCTGCGGCAATCGTGGGTTGATTTGGCGGATAAGTATATGGTAGAGGCACGGTGGTTC 1134
Db      408 TyrLeuArgLysAlaTrpValAspLeuValGluSerTyrLeuIleGluAlaLysTrpTyr 427

Qy     1135 TACGGCGGGCACAAACCAAGTTTGGAAGAGTATTTGGAGAACTCATGGCAGTCGATAAGT 1194
Db      428 TyrMetGlyHisLysProSerLeuGluGluTyrMetLysAsnSerTrpIleSerIleGly 447

Qy     1195 GGGCCCTGTATGTTAACGCACATATTCTTCCGAGTAACAGATTCGTTCACAAAGGAGACC 1254
Db      448 GlyIleProIleLeuSerHisLeuPhePheArgLeuThrAspSerIleGluGluGluAsp 467

Qy     1255 GTCGACAGTTTGTACAAATACCACGATTTAGTTCGTTGGTCATCCTTCGTTCTGCGGCTT 1314
Db      468 AlaGluSerMetHisLysTyrHisAspIleValArgAlaSerCysThrIleLeuArgLeu 487

Qy     1315 GCTGATGATTTGGGAACCTCGGTGGAAGAGGTGAGCAGAGGGGATGTGCCGAAATCACTT 1374
Db      488 AlaAspAspMetGlyThrSerLeuAspGluValGluArgGlyAspValProLysSerVal 507

Qy     1375 CAGTGCTACATGAGTGACTACAATGCATCGGAGGCGGAGGCGCGGAAGCACGTGAAATGG 1434
Db      508 GlnCysTyrMetAsnGluLysAsnAlaSerGluGluGluAlaArgGluHisValArgSer 527

Qy     1435 CTGATAGCGGAGGTGTGGAAGAAGATGAATGCGGAGAGGGTGTCGAAGGATTCTCCATTC 1494
Db      528 LeuIleAspGlnThrTrpLysMetMetAsnLysGluMetMetThr------SerSerPhe 545

Qy     1495 GGCAAAGATTTTATAGGATGTGCAGTTGATTTAGGAAGGATGGCGCAGTTGATGTAC--- 1551
Db      546 SerLysTyrPheValGlnValSerAlaAsnLeuAlaArgMetAlaGlnTrpIleTyrGln 565

Qy     1552 CATAATGGAGATGGGCACGGCACACAACACCCTATTATACATCAACAAATGACCAGAACC 1611
Db      566 HisGluSerAspGlyPheGlyMetGlnHisSerLeuValAsnLysMetLeuArgGlyLeu 585

Qy     1612 TTATTCGAGCCCTTT 1626
Db      586 LeuPheAspArgTyr 590
```

```
RESULT 13
US-09-900-797-24
; Sequence 24, Application US/09900797
; Patent No. 6645762
; GENERAL INFORMATION:
;  APPLICANT: Chappell, Joseph
;  APPLICANT:  No. 66457621, Joseph P.
;  APPLICANT:  Starks, Courtney M.
;  APPLICANT:  Manna, Kathleen R.
;  TITLE OF INVENTION: SYNTHASES
;  FILE REFERENCE: 07678-025001
;  CURRENT APPLICATION NUMBER: US/09/900,797
;  CURRENT FILING DATE:  2001-07-06
;  PRIOR APPLICATION NUMBER: US/09/398,395
;  PRIOR FILING DATE: 1999-09-17
;  PRIOR APPLICATION NUMBER: 60/130,628
;  PRIOR FILING DATE: 1999-04-22
;  PRIOR APPLICATION NUMBER: 60/150,262
;  PRIOR FILING DATE: 1999-08-23
;  NUMBER OF SEQ ID NOS: 58
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 24
;   LENGTH: 591
;   TYPE: PRT
;   ORGANISM: Salvia officinalis
US-09-900-797-24

Alignment Scores:
Pred. No.:              1.26e-157    Length:         591
Score:                  1633.50      Matches:        305
Percent Similarity:     75.96%       Conservative:   109
Best Local Similarity:  55.96%       Mismatches:     116
Query Match:            56.17%       Indels:         15
DB:                     4            Gaps:           7

US-09-938-956-6 (1-1632) x US-09-900-797-24 (1-591)

Qy        4 AGACGATCCGGAAACTACAACCCTTCTCGTTGGGATGTCAACTTCATCCAATCGCTTCTC 63
Db       57 ArgArgThrGlyGlyTyrGlnProThrLeuTrpAspPheSerThrIleGlnLeuPheAsp 76

Qy       64 AGTGACTATAAGGAGGACAAACACGTGATTAGGGCTTCTGAGCTGGTCACTTTGGTGAAG 123
Db       77 SerGluTyrLysGluGluLysHisLeuMetArgAlaAlaGlyMetIleAlaGlnValAsn 96

Qy      124 ATGGAACTGGAGAAAGAAACGGATCAAATTCGACAACTTGAGTTGATCGATGACTTGCAG 183
Db       97 MetLeuLeuGlnGluGluValAspSerIleGlnArgLeuGluLeuIleAspAspLeuArg 116

Qy      184 AGGATGGGGCTGTCCGATCATTTCCAAAATGAGTTCAAAGAAATCTTGTCCTCTATATAT 243
Db      117 ArgLeuGlyIleSerCysHisPheAspArgGluIleValGluIleLeuAsnSerLysTyr 136

Qy      244 CTCGACCATCACTATTACAAGAACCCTTTTCCAAAAGAAGAAAGGGATCTCTACTCCACA 303
Db      137 ------------TyrThrAsnAsnGluIle------AspGluSerAspLeuTyrSerThr 150

Qy      304 TCTCTTGCATTTAGGCTCCTCAGAGAACATGGTTTTCAAGTCGCACAAGAGGTATTCGAT 363
Db      151 AlaLeuArgPheLysLeuLeuArgGlnTyrAspPheSerValSerGlnGluValPheAsp 170

Qy      364 AGTTTCAAGAACGAGGAGGGT---GAGTTCAAAGAAAGCCTTAGCGACGACACCAGAGGA 420
Db      171 CysPheLysAsnAspLysGlyThrAspPheLysProSerLeuValAspAspThrArgGly 190

Qy      421 TTGTTGCAACTGTATGAAGCTTCCTTTCTGTTGACGGAAGGCGAAACCACGCTCGAGTCA 480
Db      191 LeuLeuGlnLeuTyrGluAlaSerPheLeuSerAlaGlnGlyGluGluThrLeuHisLeu 210

Qy      481 GCGAGGGAATTCGCCACCAAATTTTTGGAGGAAAAAGTGAACGAGGGTGGTGTTGATGGC 540
Db      211 AlaArgAspPheAlaThrLysPheLeuHisLysArgValLeu---------ValAspLys 227

Qy      541 GAC------CTTTTAACAAGAATCGCATATTCTTTGGACATCCCTCTTCATTGGAGGATT 594
Db      228 AspIleAsnLeuLeuSerSerIleGluArgAlaLeuGluLeuProThrHisTrpArgVal 247

Qy      595 AAAAGGCCAAATGCACCTGTGTGGATCGAATGGTATAGGAAGAGGCCCGACATGAATCCA 654
Db      248 GlnMetProAsnAlaArgSerPheIleAspAlaTyrLysArgArgProAspMetAsnPro 267

Qy      655 GTAGTGTTGGAGCTTGCCATACTCGACTTAAATATTGTTCAAGCACAATTTCAAGAAGAG 714
Db      268 ThrValLeuGluLeuAlaLysLeuAspPheAsnMetValGlnAlaGlnPheGlnGluGlu 287

Qy      715 CTCAAAGAATCCTTCAGGTGGTGGAGAAATACTGGGTTTGTTGAGAAGCTGCCCTTCGCA 774
Db      288 LeuLysGluAlaSerArgTrpTrpAsnSerThrGlyLeuValHisGluLeuProPheVal 307

Qy      775 AGGGATAGACTGGTGGAATGCTACTTTTGGAATACTGGGATCATCGAGCCACGTCAGCAT 834
Db      308 ArgAspArgIleValGluCysTyrTyrTrpThrThrGlyValValGluArgArgGluHis 327

Qy      835 GCAAGTGCAAGGATAATGATGGGCAAAGTCAACGCTCTGATTACGGTGATCGATGATATT 894
Db      328 GlyTyrGluArgIleMetLeuThrLysIleAsnAlaLeuValThrThrIleAspAspVal 347

Qy      895 TATGATGTCTATGGCACCTTAGAAGAACTCGAACAATTCACTGACCTCATTCGAAGATGG 954
Db      348 PheAspIleTyrGlyThrLeuGluGluLeuGlnLeuPheThrThrAlaIleGlnArgTrp 367

Qy      955 GATATAAACTCAATCGACCAACTTCCCGATTACATGCAACTGTGCTTTCTTGCACTCAAC 1014
Db      368 AspIleGluSerMetLysGlnLeuProProTyrMetGlnIleCysTyrLeuAlaLeuPhe 387

Qy     1015 AACTTCGTCGATGATACATCGTACGATGTTATGAAGGAGAAAGGCGTCAACGTTATACCC 1074
```

```
Db     388 AsnPheValAsnGluMetAlaTyrAspThrLeuArgAspLysGlyPheAsnSerThrPro 407

Qy    1075 TACCTGCGGCAATCGTGGGTTGATTTGGCGGATAAGTATATGGTAGAGGCACGGTGGTTC 1134
Db     408 TyrLeuArgLysAlaTrpValAspLeuValGluSerTyrLeuIleGluAlaLysTrpTyr 427

Qy    1135 TACGGCGGGCACAAACCAAGTTTGGAAGAGTATTTGGAGAACTCATGGCAGTCGATAAGT 1194
Db     428 TyrMetGlyHisLysProSerLeuGluGluTyrMetLysAsnSerTrpIleSerIleGly 447

Qy    1195 GGGCCCTGTATGTTAACGCACATATTCTTCCGAGTAACAGATTCGTTCACAAAGGAGACC 1254
Db     448 GlyIleProIleLeuSerHisLeuPhePheArgLeuThrAspSerIleGluGluGluAsp 467

Qy    1255 GTCGACAGTTTGTACAAATACCACGATTTAGTTCGTTGGTCATCCTTCGTTCTGCGGCTT 1314
Db     468 AlaGluSerMetHisLysTyrHisAspIleValArgAlaSerCysThrIleLeuArgLeu 487

Qy    1315 GCTGATGATTTGGGAACCTCGGTGGAAGAGGTGAGCAGAGGGGATGTGCCGAAATCACTT 1374
Db     488 AlaAspAspMetGlyThrSerLeuAspGluValGluArgGlyAspValProLysSerVal 507

Qy    1375 CAGTGCTACATGAGTGACTACAATGCATCGGAGGCGGAGGCGCGGAAGCACGTGAAATGG 1434
Db     508 GlnCysTyrMetAsnGluLysAsnAlaSerGluGluGluAlaArgGluHisValArgSer 527

Qy    1435 CTGATAGCGGAGGTGTGGAAGAAGATGAATGCGGAGAGGGTGTCGAAGGATTCTCCATTC 1494
Db     528 LeuIleAspGlnThrTrpLysMetMetAsnLysGluMetMetThr------SerSerPhe 545

Qy    1495 GGCAAAGATTTTATAGGATGTGCAGTTGATTTAGGAAGGATGGCGCAGTTGATGTAC--- 1551
Db     546 SerLysTyrPheValGlnValSerAlaAsnLeuAlaArgMetAlaGlnTrpIleTyrGln 565

Qy    1552 CATAATGGAGATGGGCACGGCACACAACACCCTATTATACATCAACAAATGACCAGAACC 1611
Db     566 HisGluSerAspGlyPheGlyMetGlnHisSerLeuValAsnLysMetLeuArgGlyLeu 585

Qy    1612 TTATTCGAGCCCTTT 1626
Db     586 LeuPheAspArgTyr 590
```

RESULT 14
US-08-937-540-6

```
; Sequence 6, Application US/08937540
; Patent No. 5891697
;  GENERAL INFORMATION:
;    APPLICANT:  Croteau, Rodney B
;    APPLICANT:  Wise, Mitchell L
;    APPLICANT:  Savage, Thomas J
;    APPLICANT:  Katahira, Eva J
;    TITLE OF INVENTION:  Monoterpene Synthases from Common Sage
;    TITLE OF INVENTION:  (Salvia officinalis)
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  CHRISTENSEN, O'CONNOR, JOHNSON & KINDNESS
;      STREET:  1420 FIFTH AVENUE
;      CITY:  SEATTLE
;      STATE:  WASHINGTON
;      COUNTRY:  USA
;      ZIP:  98101-2347
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/937,540
;      FILING DATE:
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Shelton, Dennis K
;      REGISTRATION NUMBER:  26,997
;      REFERENCE/DOCKET NUMBER:  WSUR111254
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  206 695 1718
;      TELEFAX:  206 224 0779
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  589 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-937-540-6

Alignment Scores:
Pred. No.:              1.82e-149     Length:          589
Score:                  1553.50       Matches:         291
Percent Similarity:     72.48%        Conservative:    104
Best Local Similarity:  53.39%        Mismatches:      141
Query Match:            53.42%        Indels:          9
DB:                     2             Gaps:            3

US-09-938-956-6 (1-1632) x US-08-937-540-6 (1-589)

Qy       1 ATGAGACGATCCGGAAACTACAACCCTTCTCGTTGGGATGTCAACTTCATCCAATCGCTT 60
Db      51 IleArgArgSerGlyAspTyrGlnProSerLeuTrpAspPheAsnTyrIleGlnSerLeu 70

Qy      61 CTCAGTGACTATAAGGAGGACAAACACGTGATTAGGGCTTCTGAGCTGGTCACTTTGGTG 120
Db      71 AsnThrProTyrLysGluGlnArgHisPheAsnArgGlnAlaGluLeuIleMetGlnVal 90

Qy     121 AAGATGGAACTGGAGAAAGAAACGGATCAAATTCGACAACTTGAGTTGATCGATGACTTG 180
Db      91 ArgMetLeuLeuLysValLysMetGluAlaIleGlnGlnLeuGluLeuIleAspAspLeu 110

Qy     181 CAGAGGATGGGGCTGTCCGATCATTTCCAAAATGAGTTCAAAGAAATCTTGTCCTCTATA 240
Db     111 GlnTyrLeuGlyLeuSerTyrPhePheGlnAspGluIleLysGlnIleLeuSerSerIle 130

Qy     241 TATCTCGACCATCACTATTACAAGAACCCTTTTCCAAAAGAAGAAAGGGATCTCTACTCC 300
Db     131 HisAsnGluProArgTyrPheHisAsn-----------------AsnAspLeuTyrPhe 144

Qy     301 ACATCTCTTGCATTTAGGCTCCTCAGAGAACATGGTTTTCAAGTCGCACAAGAGGTATTC 360
Db     145 ThrAlaLeuGlyPheArgIleLeuArgGlnHisGlyPheAsnValSerGluAspValPhe 164

Qy     361 GATAGTTTCAAGAACGAGGAG---GGTGAGTTCAAAGAAAGCCTTAGCGACGACACCAGA 417
Db     165 AspCysPheLysIleGluLysCysSerAspPheAsnAlaAsnLeuAlaGlnAspThrLys 184

Qy     418 GGATTGTTGCAACTGTATGAAGCTTCCTTTCTGTTGACGGAAGGCGAAACCACGCTCGAG 477
Db     185 GlyMetLeuGlnLeuTyrGluAlaSerPheLeuLeuArgGluGlyGluAspThrLeuGlu 204

Qy     478 TCAGCGAGGGAATTCGCCACCAAATTTTTGGAGGAAAAAGTGAACGAGGGTGGT------ 531
Db     205 LeuAlaArgArgPheSerThrArgSerLeuArgGluLysPheAspGluGlyGlyAspGlu 224

Qy     532 GTTGATGGCGACCTTTTAACAAGAATCGCATATTCTTTGGACATCCCTCTTCATTGGAGG 591
Db     225 IleAspGluAspLeuSerSerTrpIleArgHisSerLeuAspLeuProLeuHisTrpArg 244

Qy     592 ATTAAAAGGCCAAATGCACCTGTGTGGATCGAATGGTATAGGAAGAGGCCCGACATGAAT 651
Db     245 ValGlnGlyLeuGluAlaArgTrpPheLeuAspAlaTyrAlaArgArgProAspMetAsn 264

Qy     652 CCAGTAGTGTTGGAGCTTGCCATACTCGACTTAAATATTGTTCAAGCACAATTTCAAGAA 711
Db     265 ProLeuIlePheLysLeuAlaLysLeuAsnPheAsnIleValGlnAlaThrTyrGlnGlu 284

Qy     712 GAGCTCAAAGAATCCTTCAGGTGGTGGAGAAATACTGGGTTTGTTGAGAAGCTGCCCTTC 771
```

```
Db        285 GluLeuLysAspIleSerArgTrpTrpAsnSerSerCysLeuAlaGluLysLeuProPhe 304

Qy        772 GCAAGGGATAGACTGGTGGAATGCTACTTTTGGAATACTGGGATCATCGAGCCACGTCAG 831
              |||||||||:::|||||||||:::||||||               ||||||   |||
Db        305 ValArgAspArgIleValGluCysPhePheTrpAlaIleAlaAlaPheGluProHisGln 324

Qy        832 CATGCAAGTGCAAGGATAATGATGGGCAAAGTCAACGCTCTGATTACGGTGATCGATGAT 891
              ::::::      |||   |||         :::         ||||||:::|||||||||
Db        325 TyrSerTyrGlnArgLysMetAlaAlaValIleIleThrPheIleThrIleIleAspAsp 344

Qy        892 ATTTATGATGTCTATGGCACCTTAGAAGAACTCGAACAATTCACTGACCTCATTCGAAGA 951
              :::||||||||||||||||||:::||||||||||||      ||||||:::|||||||||
Db        345 ValTyrAspValTyrGlyThrIleGluGluLeuGluLeuLeuThrAspMetIleArgArg 364

Qy        952 TGGGATATAAACTCAATCGACCAACTTCCCGATTACATGCAACTGTGCTTTCTTGCACTC 1011
              ||||||      ||||||   |||||||||   |||||||||:::|||:::|||||||||
Db        365 TrpAspAsnLysSerIleSerGlnLeuProTyrTyrMetGlnValCysTyrLeuAlaLeu 384

Qy       1012 AACAACTTCGTCGATGATACATCGTACGATGTTATGAAGGAGAAAGGCGTCAACGTTATA 1071
              |||||||||   :::   :::||||||::::::|||::::::         |||   |||
Db        385 TyrAsnPheValSerGluArgAlaTyrAspIleLeuLysAspGlnHisPheAsnSerIle 404

Qy       1072 CCCTACCTGCGGCAATCGTGGGTTGATTTGGCGGATAAGTATATGGTAGAGGCACGGTGG 1131
              |||||||||::::::|||||||||   |||   :::   |||:::   ||||||   |||
Db        405 ProTyrLeuGlnArgSerTrpValSerLeuValGluGlyTyrLeuLysGluAlaTyrTrp 424

Qy       1132 TTCTACGGCGGGCACAAACCAAGTTTGGAAGAGTATTTGGAGAACTCATGGCAGTCGATA 1191
              :::|||   |||:::||||||||||||||||||||||||   |||:::      ||||||
Db        425 TyrTyrAsnGlyTyrLysProSerLeuGluGluTyrLeuAsnAsnAlaLysIleSerIle 444

Qy       1192 AGTGGGCCCTGTATGTTAACGCACATATTCTTCCGAGTAACAGATTCGTTCACAAAGGAG 1251
              |||   |||   :::::::::   ::::::|||   :::   :::|||      :::
Db        445 SerAlaProThrIleIleSerGlnLeuTyrPheThrLeuAlaAsnSerIleAspGluThr 464

Qy       1252 ACCGTCGACAGTTTGTACAAATACCACGATTTAGTTCGTTGGTCATCCTTCGTTCTGCGG 1311
              ::::::|||||||||:::||||||:::::::::         |||      :::||||||
Db        465 AlaIleGluSerLeuTyrGlnTyrHisAsnIleLeuTyrLeuSerGlyThrIleLeuArg 484

Qy       1312 CTTGCTGATGATTTGGGAACCTCGGTGGAAGAGGTGAGCAGAGGGGATGTGCCGAAATCA 1371
              ||||||||||||||||||||||||         |||:::   ||||||||||||||||:::
Db        485 LeuAlaAspAspLeuGlyThrSerGlnHisGluLeuGluArgGlyAspValProLysAla 504

Qy       1372 CTTCAGTGCTACATGAGTGACTACAATGCATCGGAGGCGGAGGCGCGGAAGCACGTGAAA 1431
              :::||||||||||||:::|||   ||||||||||||   ||||||   :::|||||||||
Db        505 IleGlnCysTyrMetAsnAspThrAsnAlaSerGluArgGluAlaValGluHisValLys 524

Qy       1432 TGGCTGATAGCGGAGGTGTGGAAGAAGATGAATGCGGAGAGGGTGTCGAAGGATTCTCCA 1491
              :::||||||   |||   ||||||:::||||||               :::   |||   |||
Db        525 PheLeuIleArgGluAlaTrpLysGluMetAsnThrValThrThrAlaSerAspCysPro 544

Qy       1492 TTCGGCAAAGATTTTATAGGATGTGCAGTTGATTTAGGAAGGATGGCGCAGTTGATGTAC 1551
              |||      |||   :::      |||   :::|||   |||   ||||||   :::|||
Db        545 PheThrAspAspLeuValAlaAlaAlaAlaAsnLeuAlaArgAlaAlaGlnPheIleTyr 564

Qy       1552 CATAATGGAGATGGGCACGGCACACAACACCCTATTATACATCAACAAATGACCAGAACC 1611
                 :::|||||||||||||||   ||||||      ||||||||||||||||
Db        565 LeuAspGlyAspGlyHisGlyValGlnHisSerGluIleHisGlnGlnMetGlyGlyLeu 584

Qy       1612 TTATTCGAGCCCTTT 1626
              ||||||:::|||:::
Db        585 LeuPheGlnProTyr 589

RESULT 15
US-09-398-395A-54
; Sequence 54, Application US/09398395A
; Patent No. 6468772
; GENERAL INFORMATION:
;  APPLICANT: Chappell, Joseph
;  APPLICANT:  No. 64687721, Joseph P.
;  APPLICANT:  Starks, Courtney M.
;  APPLICANT:  Manna, Kathleen R.
;  TITLE OF INVENTION: SYNTHASES
;  FILE REFERENCE: 07678-025001
;  CURRENT APPLICATION NUMBER: US/09/398,395A
;  CURRENT FILING DATE:  1999-09-17
;  PRIOR APPLICATION NUMBER: 60/100,993
;  PRIOR FILING DATE: 1998-09-18
;  PRIOR APPLICATION NUMBER: 60/130,628
;  PRIOR FILING DATE: 1999-04-22
;  PRIOR APPLICATION NUMBER: 60/150,262
;  PRIOR FILING DATE: 1999-08-23
;  NUMBER OF SEQ ID NOS: 58
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 54
;   LENGTH: 590
;   TYPE: PRT
;   ORGANISM: Salvia officinalis
US-09-398-395A-54

Alignment Scores:
Pred. No.:              1.82e-149      Length:         590
Score:                  1553.50        Matches:        291
Percent Similarity:     72.48%         Conservative:   104
Best Local Similarity:  53.39%         Mismatches:     141
Query Match:            53.42%         Indels:         9
DB:                     4              Gaps:           3

US-09-938-956-6 (1-1632) x US-09-398-395A-54 (1-590)

Qy          1 ATGAGACGATCCGGAAACTACAACCCTTCTCGTTGGGATGTCAACTTCATCCAATCGCTT 60
              :::||||||||||||:::|||   ||||||   ||||||   |||:::||||||||||||
Db         51 IleArgArgSerGlyAspTyrGlnProSerLeuTrpAspPheAsnTyrIleGlnSerLeu 70

Qy         61 CTCAGTGACTATAAGGAGGACAAACACGTGATTAGGGCTTCTGAGCTGGTCACTTTGGTG 120
                 :::   |||||||||   :::|||         |||   :::|||||||:::      |||
Db         71 AsnThrProTyrLysGluGlnArgHisPheAsnArgGlnAlaGluLeuIleMetGlnVal 90

Qy        121 AAGATGGAACTGGAGAAAGAAACGGATCAAATTCGACAACTTGAGTTGATCGATGACTTG 180
              :::|||   |||:::   :::   :::   |||:::|||||||||||||||||||||||||
Db         91 ArgMetLeuLeuLysValLysMetGluAlaIleGlnGlnLeuGluLeuIleAspAspLeu 110

Qy        181 CAGAGGATGGGGCTGTCCGATCATTTCCAAAATGAGTTCAAAGAAATCTTGTCCTCTATA 240
              |||   :::|||||||||         |||||||:::|||   |||:::|||||||||||||||
Db        111 GlnTyrLeuGlyLeuSerTyrPhePheGlnAspGluIleLysGlnIleLeuSerSerIle 130

Qy        241 TATCTCGACCATCACTATTACAAGAACCCTTTTCCAAAAGAAGAAAGGGATCTCTACTCC 300
              :::   :::         |||:::   |||                        |||||||||
Db        131 HisAsnGluProArgTyrPheHisAsn------------------AsnAspLeuTyrPhe 144

Qy        301 ACATCTCTTGCATTTAGGCTCCTCAGAGAACATGGTTTTCAAGTCGCACAAGAGGTATTC 360
              |||:::|||   ||||||:::||||||:::|||||||||      |||:::::::::||||||
Db        145 ThrAlaLeuGlyPheArgIleLeuArgGlnHisGlyPheAsnValSerGluAspValPhe 164

Qy        361 GATAGTTTCAAGAACGAGGAG---GGTGAGTTCAAAGAAAGCCTTAGCGACGACACCAGA 417
              |||   ||||||   |||:::         :::|||      :::|||:::   ||||||:::
Db        165 AspCysPheLysIleGluLysCysSerAspPheAsnAlaAsnLeuAlaGlnAspThrLys 184

Qy        418 GGATTGTTGCAACTGTATGAAGCTTCCTTTCTGTTGACGGAAGGCGAAACCACGCTCGAG 477
              |||:::|||||||||||||||||||||||||||      |||||||||   |||||||||
Db        185 GlyMetLeuGlnLeuTyrGluAlaSerPheLeuLeuArgGluGlyGluAspThrLeuGlu 204

Qy        478 TCAGCGAGGGAATTCGCCACCAAATTTTGGAGGAAAAAGTGAACGAGGGTGGT------ 531
                 ||||||   |||:::|||:::   |||   ||||||   :::|||||||||
Db        205 LeuAlaArgArgPheSerThrArgSerLeuArgGluLysPheAspGluGlyGlyAspGlu 224

Qy        532 GTTGATGGCGACCTTTTAACAAGAATCGCATATTCTTTGGACATCCCTCTTCATTGGAGG 591
              :::|||   ||||||   :::   |||   :::|||||||||:::|||||||||||||||
Db        225 IleAspGluAspLeuSerSerTrpIleArgHisSerLeuAspLeuProLeuHisTrpArg 244

Qy        592 ATTAAAAGGCCAAATGCACCTGTGTGGATCGAATGGTATAGGAAGAGGCCCGACATGAAT 651
              ::::::      |||      :::::::::   |||   :::|||||||||||||||
Db        245 ValGlnGlyLeuGluAlaArgTrpPheLeuAspAlaTyrAlaArgArgProAspMetAsn 264

Qy        652 CCAGTAGTGTTGGAGCTTGCCATACTCGACTTAAATATTGTTCAAGCACAATTTCAAGAA 711
              |||:::::::   :::||||||   |||:::   |||||||||||||||   :::||||||
Db        265 ProLeuIlePheLysLeuAlaLysLeuAsnPheAsnIleValGlnAlaThrTyrGlnGlu 284
```

```
Qy        712 GAGCTCAAAGAATCCTTCAGGTGGTGGAGAAATACTGGGTTTGTTGAGAAGCTGCCCTTC 771
              |||||||||:::      |||||||||   ::::::       ||||||||||||||
Db        285 GluLeuLysAspIleSerArgTrpTrpAsnSerSerCysLeuAlaGluLysLeuProPhe 304

Qy        772 GCAAGGGATAGACTGGTGGAATGCTACTTTTGGAATACTGGGATCATCGAGCCACGTCAG 831
                 ||||||||:::|||||||||:::||||||             ||||||   |||
Db        305 ValArgAspArgIleValGluCysPhePheTrpAlaIleAlaAlaPheGluProHisGln 324

Qy        832 CATGCAAGTGCAAGGATAATGATGGGCAAAGTCAACGCTCTGATTACGGTGATCGATGAT 891
              ::::::     |||   |||         :::         ||||||:::|||||||||
Db        325 TyrSerTyrGlnArgLysMetAlaAlaValIleIleThrPheIleThrIleIleAspAsp 344

Qy        892 ATTTATGATGTCTATGGCACCTTAGAAGAACTCGAACAATTCACTGACCTCATTCGAAGA 951
              :::||||||||||||||||||:::||||||||||||      ||||||:::|||||||||
Db        345 ValTyrAspValTyrGlyThrIleGluGluLeuGluLeuLeuThrAspMetIleArgArg 364

Qy        952 TGGGATATAAACTCAATCGACCAACTTCCCGATTACATGCAACTGTGCTTTCTTGCACTC 1011
              ||||||      ||||||   |||||||||   |||||||||:::|||:::|||||||||
Db        365 TrpAspAsnLysSerIleSerGlnLeuProTyrTyrMetGlnValCysTyrLeuAlaLeu 384

Qy       1012 AACAACTTCGTCGATGATACATCGTACGATGTTATGAAGGAGAAAGGCGTCAACGTTATA 1071
                 ||||||||   :::   :::||||||::::::|||::::::      |||   |||
Db        385 TyrAsnPheValSerGluArgAlaTyrAspIleLeuLysAspGlnHisPheAsnSerIle 404

Qy       1072 CCCTACCTGCGGCAATCGTGGGTTGATTTGGCGGATAAGTATATGGTAGAGGCACGGTGG 1131
              |||||||||::::::|||||||||   |||   :::   |||:::   ||||||   |||
Db        405 ProTyrLeuGlnArgSerTrpValSerLeuValGluGlyTyrLeuLysGluAlaTyrTrp 424

Qy       1132 TTCTACGGCGGGCACAAACCAAGTTTGGAAGAGTATTTGGAGAACTCATGGCAGTCGATA 1191
              :::|||   |||:::||||||||||||||||||||||||   |||:::      ||||||
Db        425 TyrTyrAsnGlyTyrLysProSerLeuGluGluTyrLeuAsnAsnAlaLysIleSerIle 444

Qy       1192 AGTGGGCCCTGTATGTTAACGCACATATTCTTCCGAGTAACAGATTCGTTCACAAAGGAG 1251
              |||   |||   :::::::::   ::::::|||   :::   :::|||      :::
Db        445 SerAlaProThrIleIleSerGlnLeuTyrPheThrLeuAlaAsnSerIleAspGluThr 464

Qy       1252 ACCGTCGACAGTTTGTACAAATACCACGATTTAGTTCGTTGGTCATCCTTCGTTCTGCGG 1311
                 ::::::|||||||||:::||||||:::::::::      |||      :::||||||
Db        465 AlaIleGluSerLeuTyrGlnTyrHisAsnIleLeuTyrLeuSerGlyThrIleLeuArg 484

Qy       1312 CTTGCTGATGATTTGGGAACCTCGGTGGAAGAGGTGAGCAGAGGGGATGTGCCGAAATCA 1371
              ||||||||||||||||||||||||      |||:::   ||||||||||||||||||:::
Db        485 LeuAlaAspAspLeuGlyThrSerGlnHisGluLeuGluArgGlyAspValProLysAla 504

Qy       1372 CTTCAGTGCTACATGAGTGACTACAATGCATCGGAGGCGGAGGCGCGGAAGCACGTGAAA 1431
              :::||||||||||||:::|||   ||||||||||||   ||||||   :::|||||||||
Db        505 IleGlnCysTyrMetAsnAspThrAsnAlaSerGluArgGluAlaValGluHisValLys 524

Qy       1432 TGGCTGATAGCGGAGGTGTGGAAGAAGATGAATGCGGAGAGGGTGTCGAAGGATTCTCCA 1491
              :::||||||   |||   ||||||:::||||||             :::   |||   |||
Db        525 PheLeuIleArgGluAlaTrpLysGluMetAsnThrValThrThrAlaSerAspCysPro 544

Qy       1492 TTCGGCAAAGATTTTATAGGATGTGCAGTTGATTTAGGAAGGATGGCGCAGTTGATGTAC 1551
              |||      |||   :::      |||   :::|||   |||   ||||||   :::|||
Db        545 PheThrAspAspLeuValAlaAlaAlaAlaAsnLeuAlaArgAlaAlaGlnPheIleTyr 564

Qy       1552 CATAATGGAGATGGGCACGGCACACAACACCCTATTATACATCAACAAATGACCAGAACC 1611
                 :::||||||||||||||||   ||||||      ||||||||||||||
Db        565 LeuAspGlyAspGlyHisGlyValGlnHisSerGluIleHisGlnGlnMetGlyGlyLeu 584

Qy       1612 TTATTCGAGCCCTTT 1626
              ||||||:::|||:::
Db        585 LeuPheGlnProTyr 589
```

Search completed: February 25, 2004, 06:00:37
Job time : 64 secs

```
OM nucleic - protein search, using frame_plus_n2p model

Run on:         February 25, 2004, 05:49:18 ; Search time 61 Seconds
                                          (without alignments)
                                          11298.430 Million cell updates/sec

Title:          US-09-938-956-6
Perfect score:  2908
Sequence:       1 atgagacgatccggaaacta..........tattcgagccctttgcatga 1632

Scoring table:  BLOSUM62
                Xgapop 10.0 , Xgapext  0.5
                Ygapop 10.0 , Ygapext  0.5
                Fgapop  6.0 , Fgapext  7.0
                Delop   6.0 , Delext   7.0

Searched:       809742 seqs, 211153259 residues

Total number of hits satisfying chosen parameters:      1619484

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US09938956/runat_24022004_144705_3180/app_query.fasta_1.1799
-DB=Published_Applications_AA -QFMT=fastan -SUFFIX=rapb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09938956_@CGN_1_1_53_@runat_24022004_144705_3180
-NCPU=6 -ICPU=3 -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :      Published_Applications_AA:*
                1:  /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
                2:  /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
                3:  /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
                4:  /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
                5:  /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
                6:  /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
                7:  /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
                8:  /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
                9:  /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
                10:  /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
                11:  /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
                12:  /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
                13:  /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
                14:  /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
                15:  /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
                16:  /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
                17:  /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
                18:  /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2881	99.1	543	9	US-09-938-956-7	Sequence 7, Appli
2	2876	98.9	599	9	US-09-887-586A-22	Sequence 22, Appl
3	2876	98.9	599	9	US-09-903-012-22	Sequence 22, Appl
4	2876	98.9	599	10	US-09-900-797-22	Sequence 22, Appl
5	1633.5	56.2	591	9	US-09-887-586A-24	Sequence 24, Appl
6	1633.5	56.2	591	9	US-09-903-012-24	Sequence 24, Appl
7	1633.5	56.2	591	10	US-09-900-797-24	Sequence 24, Appl
8	1553.5	53.4	590	9	US-09-887-586A-54	Sequence 54, Appl
9	1553.5	53.4	590	9	US-09-903-012-54	Sequence 54, Appl
10	1553.5	53.4	590	10	US-09-900-797-54	Sequence 54, Appl
11	1468	50.5	598	9	US-09-887-586A-26	Sequence 26, Appl
12	1468	50.5	598	9	US-09-903-012-26	Sequence 26, Appl
13	1468	50.5	598	10	US-09-900-797-26	Sequence 26, Appl
14	836	28.7	556	9	US-09-887-586A-32	Sequence 32, Appl
15	836	28.7	556	9	US-09-903-012-32	Sequence 32, Appl
16	836	28.7	556	10	US-09-900-797-32	Sequence 32, Appl
17	806.5	27.7	548	9	US-09-887-586A-2	Sequence 2, Appli
18	806.5	27.7	548	9	US-09-903-012-2	Sequence 2, Appli
19	806.5	27.7	548	10	US-09-900-797-2	Sequence 2, Appli
20	804.5	27.7	548	9	US-09-887-586A-8	Sequence 8, Appli
21	804.5	27.7	548	9	US-09-887-586A-12	Sequence 12, Appl
22	804.5	27.7	548	9	US-09-903-012-8	Sequence 8, Appli
23	804.5	27.7	548	9	US-09-903-012-12	Sequence 12, Appl
24	804.5	27.7	548	10	US-09-900-797-8	Sequence 8, Appli
25	804.5	27.7	548	10	US-09-900-797-12	Sequence 12, Appl
26	803.5	27.6	548	9	US-09-887-586A-6	Sequence 6, Appli
27	803.5	27.6	548	9	US-09-903-012-6	Sequence 6, Appli
28	803.5	27.6	548	10	US-09-900-797-6	Sequence 6, Appli
29	793.5	27.3	548	9	US-09-887-586A-4	Sequence 4, Appli
30	793.5	27.3	548	9	US-09-903-012-4	Sequence 4, Appli
31	793.5	27.3	548	10	US-09-900-797-4	Sequence 4, Appli
32	792.5	27.3	548	9	US-09-887-586A-10	Sequence 10, Appl
33	792.5	27.3	548	9	US-09-903-012-10	Sequence 10, Appl
34	792.5	27.3	548	10	US-09-900-797-10	Sequence 10, Appl
35	734	25.2	548	9	US-09-887-586A-52	Sequence 52, Appl
36	734	25.2	548	9	US-09-903-012-52	Sequence 52, Appl
37	734	25.2	548	10	US-09-900-797-52	Sequence 52, Appl
38	715.5	24.6	618	14	US-10-025-145A-65	Sequence 65, Appl
39	714	24.6	628	9	US-09-887-586A-20	Sequence 20, Appl
40	714	24.6	628	9	US-09-903-012-20	Sequence 20, Appl
41	714	24.6	628	10	US-09-900-797-20	Sequence 20, Appl
42	714	24.6	628	13	US-10-041-007-22	Sequence 22, Appl
43	714	24.6	628	14	US-10-025-145A-4	Sequence 4, Appli
44	704.5	24.2	551	15	US-10-411-066-40	Sequence 40, Appl
45	700	24.1	630	14	US-10-025-145A-67	Sequence 67, Appl

ALIGNMENTS

```
RESULT 1
US-09-938-956-7
; Sequence 7, Application US/09938956
; Patent No. US20020142408A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, Siqun
;  APPLICANT:  Dicosimo, Deana J.
;  APPLICANT:  Koffas, Mattheos
;  APPLICANT:  Odom , J. Martin
;  TITLE OF INVENTION: Production of Monoterpene
;  FILE REFERENCE: CL1809 US NA
;  CURRENT APPLICATION NUMBER: US/09/938,956
;  CURRENT FILING DATE:  2001-08-24
;  PRIOR APPLICATION NUMBER: 60/229,907
;  PRIOR FILING DATE: 2000-09-0
;  PRIOR APPLICATION NUMBER: 60/229,858
;  PRIOR FILING DATE: 2000-09-01
;  NUMBER OF SEQ ID NOS: 7
;  SOFTWARE: Microsoft Office 97
; SEQ ID NO 7
;   LENGTH: 543
;   TYPE: PRT
;   ORGANISM: Mentha spicata
US-09-938-956-7
```

```
Alignment Scores:
Pred. No.:                 1.36e-260     Length:          543
Score:                     2881.00       Matches:         543
Percent Similarity:        100.00%       Conservative:    0
Best Local Similarity:     100.00%       Mismatches:      0
Query Match:               99.07%        Indels:          0
DB:                        9             Gaps:            0

US-09-938-956-6 (1-1632) x US-09-938-956-7 (1-543)

Qy          1 ATGAGACGATCCGGAAACTACAACCCTTCTCGTTGGGATGTCAACTTCATCCAATCGCTT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MetArgArgSerGlyAsnTyrAsnProSerArgTrpAspValAsnPheIleGlnSerLeu 20

Qy         61 CTCAGTGACTATAAGGAGGACAAACACGTGATTAGGGCTTCTGAGCTGGTCACTTTGGTG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 LeuSerAspTyrLysGluAspLysHisValIleArgAlaSerGluLeuValThrLeuVal 40

Qy        121 AAGATGGAACTGGAGAAAGAAACGGATCAAATTCGACAACTTGAGTTGATCGATGACTTG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 LysMetGluLeuGluLysGluThrAspGlnIleArgGlnLeuGluLeuIleAspAspLeu 60

Qy        181 CAGAGGATGGGGCTGTCCGATCATTTCCAAAATGAGTTCAAAGAAATCTTGTCCTCTATA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GlnArgMetGlyLeuSerAspHisPheGlnAsnGluPheLysGluIleLeuSerSerIle 80

Qy        241 TATCTCGACCATCACTATTACAAGAACCCTTTTCCAAAAGAAGAAAGGGATCTCTACTCC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 TyrLeuAspHisHisTyrTyrLysAsnProPheProLysGluGluArgAspLeuTyrSer 100

Qy        301 ACATCTCTTGCATTTAGGCTCCTCAGAGAACATGGTTTTCAAGTCGCACAAGAGGTATTC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 ThrSerLeuAlaPheArgLeuLeuArgGluHisGlyPheGlnValAlaGlnGluValPhe 120

Qy        361 GATAGTTTCAAGAACGAGGAGGGTGAGTTCAAAGAAAGCCTTAGCGACGACACCAGAGGA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AspSerPheLysAsnGluGluGlyGluPheLysGluSerLeuSerAspAspThrArgGly 140

Qy        421 TTGTTGCAACTGTATGAAGCTTCCTTTCTGTTGACGGAAGGCGAAACCACGCTCGAGTCA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 LeuLeuGlnLeuTyrGluAlaSerPheLeuLeuThrGluGlyGluThrThrLeuGluSer 160

Qy        481 GCGAGGGAATTCGCCACCAAATTTTTGGAGGAAAAAGTGAACGAGGGTGGTGTTGATGGC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 AlaArgGluPheAlaThrLysPheLeuGluGluLysValAsnGluGlyGlyValAspGly 180

Qy        541 GACCTTTTAACAAGAATCGCATATTCTTTGGACATCCCTCTTCATTGGAGGATTAAAAGG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AspLeuLeuThrArgIleAlaTyrSerLeuAspIleProLeuHisTrpArgIleLysArg 200

Qy        601 CCAAATGCACCTGTGTGGATCGAATGGTATAGGAAGAGGCCCGACATGAATCCAGTAGTG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 ProAsnAlaProValTrpIleGluTrpTyrArgLysArgProAspMetAsnProValVal 220

Qy        661 TTGGAGCTTGCCATACTCGACTTAAATATTGTTCAAGCACAATTTCAAGAAGAGCTCAAA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        221 LeuGluLeuAlaIleLeuAspLeuAsnIleValGlnAlaGlnPheGlnGluGluLeuLys 240

Qy        721 GAATCCTTCAGGTGGTGGAGAAATACTGGGTTTGTTGAGAAGCTGCCCTTCGCAAGGGAT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GluSerPheArgTrpTrpArgAsnThrGlyPheValGluLysLeuProPheAlaArgAsp 260

Qy        781 AGACTGGTGGAATGCTACTTTTGGAATACTGGGATCATCGAGCCACGTCAGCATGCAAGT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 ArgLeuValGluCysTyrPheTrpAsnThrGlyIleIleGluProArgGlnHisAlaSer 280

Qy        841 GCAAGGATAATGATGGGCAAAGTCAACGCTCTGATTACGGTGATCGATGATATTTATGAT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 AlaArgIleMetMetGlyLysValAsnAlaLeuIleThrValIleAspAspIleTyrAsp 300

Qy        901 GTCTATGGCACCTTAGAAGAACTCGAACAATTCACTGACCTCATTCGAAGATGGGATATA 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ValTyrGlyThrLeuGluGluLeuGluGlnPheThrAspLeuIleArgArgTrpAspIle 320

Qy        961 AACTCAATCGACCAACTTCCCGATTACATGCAACTGTGCTTTCTTGCACTCAACAACTTC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 AsnSerIleAspGlnLeuProAspTyrMetGlnLeuCysPheLeuAlaLeuAsnAsnPhe 340

Qy       1021 GTCGATGATACATCGTACGATGTTATGAAGGAGAAAGGCGTCAACGTTATACCCTACCTG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 ValAspAspThrSerTyrAspValMetLysGluLysGlyValAsnValIleProTyrLeu 360

Qy       1081 CGGCAATCGTGGGTTGATTTGGCGGATAAGTATATGGTAGAGGCACGGTGGTTCTACGGC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ArgGlnSerTrpValAspLeuAlaAspLysTyrMetValGluAlaArgTrpPheTyrGly 380

Qy       1141 GGGCACAAACCAAGTTTGGAAGAGTATTTGGAGAACTCATGGCAGTCGATAAGTGGGCCC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        381 GlyHisLysProSerLeuGluGluTyrLeuGluAsnSerTrpGlnSerIleSerGlyPro 400

Qy       1201 TGTATGTTAACGCACATATTCTTCCGAGTAACAGATTCGTTCACAAAGGAGACCGTCGAC 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        401 CysMetLeuThrHisIlePhePheArgValThrAspSerPheThrLysGluThrValAsp 420

Qy       1261 AGTTTGTACAAATACCACGATTTAGTTCGTTGGTCATCCTTCGTTCTGCGGCTTGCTGAT 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 SerLeuTyrLysTyrHisAspLeuValArgTrpSerSerPheValLeuArgLeuAlaAsp 440

Qy       1321 GATTTGGGAACCTCGGTGGAAGAGGTGAGCAGAGGGGATGTGCCGAAATCACTTCAGTGC 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        441 AspLeuGlyThrSerValGluGluValSerArgGlyAspValProLysSerLeuGlnCys 460

Qy       1381 TACATGAGTGACTACAATGCATCGGAGGCGGAGGCGCGGAAGCACGTGAAATGGCTGATA 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        461 TyrMetSerAspTyrAsnAlaSerGluAlaGluAlaArgLysHisValLysTrpLeuIle 480

Qy       1441 GCGGAGGTGTGGAAGAAGATGAATGCGGAGAGGGTGTCGAAGGATTCTCCATTCGGCAAA 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AlaGluValTrpLysLysMetAsnAlaGluArgValSerLysAspSerProPheGlyLys 500

Qy       1501 GATTTTATAGGATGTGCAGTTGATTTAGGAAGGATGGCGCAGTTGATGTACCATAATGGA 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        501 AspPheIleGlyCysAlaValAspLeuGlyArgMetAlaGlnLeuMetTyrHisAsnGly 520

Qy       1561 GATGGGCACGGCACACAACACCCTATTATACATCAACAAATGACCAGAACCTTATTCGAG 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        521 AspGlyHisGlyThrGlnHisProIleIleHisGlnGlnMetThrArgThrLeuPheGlu 540

Qy       1621 CCCTTTGCA 1629
              |||||||||
Db        541 ProPheAla 543

RESULT 2
US-09-887-586A-22
; Sequence 22, Application US/09887586A
; Patent No. US20020094556A1
; GENERAL INFORMATION:
;  APPLICANT: Chappell, Joseph
;  APPLICANT:  No. US20020094556A11, Joseph P.
;  APPLICANT:  Starks, Courtney M.
;  APPLICANT:  Manna, Kathleen R.
;  TITLE OF INVENTION: SYNTHASES
;  FILE REFERENCE: 07678-025001
;  CURRENT APPLICATION NUMBER: US/09/887,586A
;  CURRENT FILING DATE:  2001-06-22
;  PRIOR APPLICATION NUMBER: 09/398,395
;  PRIOR FILING DATE: 1999-09-17
;  PRIOR APPLICATION NUMBER: 60/130,628
;  PRIOR FILING DATE: 1999-04-22
;  PRIOR APPLICATION NUMBER: 60/150,262
;  PRIOR FILING DATE: 1999-08-23
;  NUMBER OF SEQ ID NOS: 58
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 22
;   LENGTH: 599
```

```
;   TYPE: PRT
;   ORGANISM: Mentha spicata
US-09-887-586A-22

Alignment Scores:
Pred. No.:              4.17e-260     Length:         599
Score:                  2876.00       Matches:        542
Percent Similarity:     100.00%       Conservative:   0
Best Local Similarity:  100.00%       Mismatches:     0
Query Match:            98.90%        Indels:         0
DB:                     9             Gaps:           0

US-09-938-956-6 (1-1632) x US-09-887-586A-22 (1-599)

Qy          4 AGACGATCCGGAAACTACAACCCTTCTCGTTGGGATGTCAACTTCATCCAATCGCTTCTC 63
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         58 ArgArgSerGlyAsnTyrAsnProSerArgTrpAspValAsnPheIleGlnSerLeuLeu 77

Qy         64 AGTGACTATAAGGAGGACAAACACGTGATTAGGGCTTCTGAGCTGGTCACTTTGGTGAAG 123
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         78 SerAspTyrLysGluAspLysHisValIleArgAlaSerGluLeuValThrLeuValLys 97

Qy        124 ATGGAACTGGAGAAAGAAACGGATCAAATTCGACAACTTGAGTTGATCGATGACTTGCAG 183
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         98 MetGluLeuGluLysGluThrAspGlnIleArgGlnLeuGluLeuIleAspAspLeuGln 117

Qy        184 AGGATGGGGCTGTCCGATCATTTCCAAAATGAGTTCAAAGAAATCTTGTCCTCTATATAT 243
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        118 ArgMetGlyLeuSerAspHisPheGlnAsnGluPheLysGluIleLeuSerSerIleTyr 137

Qy        244 CTCGACCATCACTATTACAAGAACCCTTTTCCAAAAGAAGAAAGGGATCTCTACTCCACA 303
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        138 LeuAspHisHisTyrTyrLysAsnProPheProLysGluGluArgAspLeuTyrSerThr 157

Qy        304 TCTCTTGCATTTAGGCTCCTCAGAGAACATGGTTTTCAAGTCGCACAAGAGGTATTCGAT 363
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        158 SerLeuAlaPheArgLeuLeuArgGluHisGlyPheGlnValAlaGlnGluValPheAsp 177

Qy        364 AGTTTCAAGAACGAGGAGGGTGAGTTCAAAGAAAGCCTTAGCGACGACACCAGAGGATTG 423
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        178 SerPheLysAsnGluGluGlyGluPheLysGluSerLeuSerAspAspThrArgGlyLeu 197

Qy        424 TTGCAACTGTATGAAGCTTCCTTTCTGTTGACGGAAGGCGAAACCACGCTCGAGTCAGCG 483
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        198 LeuGlnLeuTyrGluAlaSerPheLeuLeuThrGluGlyGluThrThrLeuGluSerAla 217

Qy        484 AGGGAATTCGCCACCAAATTTTTGGAGGAAAAAGTGAACGAGGGTGGTGTTGATGGCGAC 543
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        218 ArgGluPheAlaThrLysPheLeuGluGluLysValAsnGluGlyGlyValAspGlyAsp 237

Qy        544 CTTTTAACAAGAATCGCATATTCTTTGGACATCCCTCTTCATTGGAGGATTAAAAGGCCA 603
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        238 LeuLeuThrArgIleAlaTyrSerLeuAspIleProLeuHisTrpArgIleLysArgPro 257

Qy        604 AATGCACCTGTGTGGATCGAATGGTATAGGAAGAGGCCCGACATGAATCCAGTAGTGTTG 663
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        258 AsnAlaProValTrpIleGluTrpTyrArgLysArgProAspMetAsnProValValLeu 277

Qy        664 GAGCTTGCCATACTCGACTTAAATATTGTTCAAGCACAATTTCAAGAAGAGCTCAAAGAA 723
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        278 GluLeuAlaIleLeuAspLeuAsnIleValGlnAlaGlnPheGlnGluGluLeuLysGlu 297

Qy        724 TCCTTCAGGTGGTGGAGAAATACTGGGTTTGTTGAGAAGCTGCCCTTCGCAAGGGATAGA 783
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        298 SerPheArgTrpTrpArgAsnThrGlyPheValGluLysLeuProPheAlaArgAspArg 317

Qy        784 CTGGTGGAATGCTACTTTTGGAATACTGGGATCATCGAGCCACGTCAGCATGCAAGTGCA 843
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        318 LeuValGluCysTyrPheTrpAsnThrGlyIleIleGluProArgGlnHisAlaSerAla 337

Qy        844 AGGATAATGATGGGCAAAGTCAACGCTCTGATTACGGTGATCGATGATATTTATGATGTC 903
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        338 ArgIleMetMetGlyLysValAsnAlaLeuIleThrValIleAspAspIleTyrAspVal 357

Qy        904 TATGGCACCTTAGAAGAACTCGAACAATTCACTGACCTCATTCGAAGATGGGATATAAAC 963
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        358 TyrGlyThrLeuGluGluLeuGluGlnPheThrAspLeuIleArgArgTrpAspIleAsn 377

Qy        964 TCAATCGACCAACTTCCCGATTACATGCAACTGTGCTTTCTTGCACTCAACAACTTCGTC 1023
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        378 SerIleAspGlnLeuProAspTyrMetGlnLeuCysPheLeuAlaLeuAsnAsnPheVal 397

Qy       1024 GATGATACATCGTACGATGTTATGAAGGAGAAAGGCGTCAACGTTATACCCTACCTGCGG 1083
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        398 AspAspThrSerTyrAspValMetLysGluLysGlyValAsnValIleProTyrLeuArg 417

Qy       1084 CAATCGTGGGTTGATTTGGCGGATAAGTATATGGTAGAGGCACGGTGGTTCTACGGCGGG 1143
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        418 GlnSerTrpValAspLeuAlaAspLysTyrMetValGluAlaArgTrpPheTyrGlyGly 437

Qy       1144 CACAAACCAAGTTTGGAAGAGTATTTGGAGAACTCATGGCAGTCGATAAGTGGGCCCTGT 1203
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        438 HisLysProSerLeuGluGluTyrLeuGluAsnSerTrpGlnSerIleSerGlyProCys 457

Qy       1204 ATGTTAACGCACATATTCTTCCGAGTAACAGATTCGTTCACAAAGGAGACCGTCGACAGT 1263
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        458 MetLeuThrHisIlePhePheArgValThrAspSerPheThrLysGluThrValAspSer 477

Qy       1264 TTGTACAAATACCACGATTTAGTTCGTTGGTCATCCTTCGTTCTGCGGCTTGCTGATGAT 1323
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        478 LeuTyrLysTyrHisAspLeuValArgTrpSerSerPheValLeuArgLeuAlaAspAsp 497

Qy       1324 TTGGGAACCTCGGTGGAAGAGGTGAGCAGAGGGGATGTGCCGAAATCACTTCAGTGCTAC 1383
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        498 LeuGlyThrSerValGluGluValSerArgGlyAspValProLysSerLeuGlnCysTyr 517

Qy       1384 ATGAGTGACTACAATGCATCGGAGGCGGAGGCGCGGAAGCACGTGAAATGGCTGATAGCG 1443
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        518 MetSerAspTyrAsnAlaSerGluAlaGluAlaArgLysHisValLysTrpLeuIleAla 537

Qy       1444 GAGGTGTGGAAGAAGATGAATGCGGAGAGGGTGTCGAAGGATTCTCCATTCGGCAAAGAT 1503
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        538 GluValTrpLysLysMetAsnAlaGluArgValSerLysAspSerProPheGlyLysAsp 557

Qy       1504 TTTATAGGATGTGCAGTTGATTTAGGAAGGATGGCGCAGTTGATGTACCATAATGGAGAT 1563
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        558 PheIleGlyCysAlaValAspLeuGlyArgMetAlaGlnLeuMetTyrHisAsnGlyAsp 577

Qy       1564 GGGCACGGCACACAACACCCTATTATACATCAACAAATGACCAGAACCTTATTCGAGCCC 1623
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        578 GlyHisGlyThrGlnHisProIleIleHisGlnGlnMetThrArgThrLeuPheGluPro 597

Qy       1624 TTTGCA 1629
              ||||||
Db        598 PheAla 599

RESULT 3
US-09-903-012-22
; Sequence 22, Application US/09903012
; Patent No. US20020094557A1
; GENERAL INFORMATION:
;  APPLICANT: Chappell, Joseph
;  APPLICANT:  No. US20020094557A11, Joseph P.
;  APPLICANT:  Starks, Courtney M.
;  APPLICANT:  Manna, Kathleen R.
;  TITLE OF INVENTION: SYNTHASES
;  FILE REFERENCE: 07678-025001
;  CURRENT APPLICATION NUMBER: US/09/903,012
;  CURRENT FILING DATE:  2001-07-11
;  PRIOR APPLICATION NUMBER: 09/398,395
;  PRIOR FILING DATE: 1999-09-17
;  PRIOR APPLICATION NUMBER: 60/100,993
;  PRIOR FILING DATE: 1998-09-18
;  PRIOR APPLICATION NUMBER: 60/130,628
;  PRIOR FILING DATE: 1999-04-22
;  PRIOR APPLICATION NUMBER: 60/150,262
```

```
;  PRIOR FILING DATE: 1999-08-23
;  NUMBER OF SEQ ID NOS: 58
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 22
;   LENGTH: 599
;   TYPE: PRT
;   ORGANISM: Mentha spicata
US-09-903-012-22

Alignment Scores:
Pred. No.:               4.17e-260      Length:         599
Score:                   2876.00        Matches:        542
Percent Similarity:      100.00%        Conservative:   0
Best Local Similarity:   100.00%        Mismatches:     0
Query Match:             98.90%         Indels:         0
DB:                      9              Gaps:           0

US-09-938-956-6 (1-1632) x US-09-903-012-22 (1-599)

Qy        4 AGACGATCCGGAAACTACAACCCTTCTCGTTGGGATGTCAACTTCATCCAATCGCTTCTC 63
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       58 ArgArgSerGlyAsnTyrAsnProSerArgTrpAspValAsnPheIleGlnSerLeuLeu 77

Qy       64 AGTGACTATAAGGAGGACAAACACGTGATTAGGGCTTCTGAGCTGGTCACTTTGGTGAAG 123
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       78 SerAspTyrLysGluAspLysHisValIleArgAlaSerGluLeuValThrLeuValLys 97

Qy      124 ATGGAACTGGAGAAAGAAACGGATCAAATTCGACAACTTGAGTTGATCGATGACTTGCAG 183
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       98 MetGluLeuGluLysGluThrAspGlnIleArgGlnLeuGluLeuIleAspAspLeuGln 117

Qy      184 AGGATGGGGCTGTCCGATCATTTCCAAAATGAGTTCAAAGAAATCTTGTCCTCTATATAT 243
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      118 ArgMetGlyLeuSerAspHisPheGlnAsnGluPheLysGluIleLeuSerSerIleTyr 137

Qy      244 CTCGACCATCACTATTACAAGAACCCTTTTCCAAAAGAAGAAAGGGATCTCTACTCCACA 303
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      138 LeuAspHisHisTyrTyrLysAsnProPheProLysGluGluArgAspLeuTyrSerThr 157

Qy      304 TCTCTTGCATTTAGGCTCCTCAGAGAACATGGTTTTCAAGTCGCACAAGAGGTATTCGAT 363
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      158 SerLeuAlaPheArgLeuLeuArgGluHisGlyPheGlnValAlaGlnGluValPheAsp 177

Qy      364 AGTTTCAAGAACGAGGAGGGTGAGTTCAAAGAAAGCCTTAGCGACGACACCAGAGGATTG 423
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      178 SerPheLysAsnGluGluGlyGluPheLysGluSerLeuSerAspAspThrArgGlyLeu 197

Qy      424 TTGCAACTGTATGAAGCTTCCTTTCTGTTGACGGAAGGCGAAACCACGCTCGAGTCAGCG 483
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      198 LeuGlnLeuTyrGluAlaSerPheLeuLeuThrGluGlyGluThrThrLeuGluSerAla 217

Qy      484 AGGGAATTCGCCACCAAATTTTTGGAGGAAAAAGTGAACGAGGGTGGTGTTGATGGCGAC 543
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      218 ArgGluPheAlaThrLysPheLeuGluGluLysValAsnGluGlyGlyValAspGlyAsp 237

Qy      544 CTTTTAACAAGAATCGCATATTCTTTGGACATCCCTCTTCATTGGAGGATTAAAAGGCCA 603
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      238 LeuLeuThrArgIleAlaTyrSerLeuAspIleProLeuHisTrpArgIleLysArgPro 257

Qy      604 AATGCACCTGTGTGGATCGAATGGTATAGGAAGAGGCCCGACATGAATCCAGTAGTGTTG 663
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      258 AsnAlaProValTrpIleGluTrpTyrArgLysArgProAspMetAsnProValValLeu 277

Qy      664 GAGCTTGCCATACTCGACTTAAATATTGTTCAAGCACAATTTCAAGAAGAGCTCAAAGAA 723
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      278 GluLeuAlaIleLeuAspLeuAsnIleValGlnAlaGlnPheGlnGluGluLeuLysGlu 297

Qy      724 TCCTTCAGGTGGTGGAGAAATACTGGGTTTGTTGAGAAGCTGCCCTTCGCAAGGGATAGA 783
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      298 SerPheArgTrpTrpArgAsnThrGlyPheValGluLysLeuProPheAlaArgAspArg 317

Qy      784 CTGGTGGAATGCTACTTTTGGAATACTGGGATCATCGAGCCACGTCAGCATGCAAGTGCA 843
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      318 LeuValGluCysTyrPheTrpAsnThrGlyIleIleGluProArgGlnHisAlaSerAla 337

Qy      844 AGGATAATGATGGGCAAAGTCAACGCTCTGATTACGGTGATCGATGATATTTATGATGTC 903
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      338 ArgIleMetMetGlyLysValAsnAlaLeuIleThrValIleAspAspIleTyrAspVal 357

Qy      904 TATGGCACCTTAGAAGAACTCGAACAATTCACTGACCTCATTCGAAGATGGGATATAAAC 963
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      358 TyrGlyThrLeuGluGluLeuGluGlnPheThrAspLeuIleArgArgTrpAspIleAsn 377

Qy      964 TCAATCGACCAACTTCCCGATTACATGCAACTGTGCTTTCTTGCACTCAACAACTTCGTC 1023
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      378 SerIleAspGlnLeuProAspTyrMetGlnLeuCysPheLeuAlaLeuAsnAsnPheVal 397

Qy     1024 GATGATACATCGTACGATGTTATGAAGGAGAAAGGCGTCAACGTTATACCCTACCTGCGG 1083
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      398 AspAspThrSerTyrAspValMetLysGluLysGlyValAsnValIleProTyrLeuArg 417

Qy     1084 CAATCGTGGGTTGATTTGGCGGATAAGTATATGGTAGAGGCACGGTGGTTCTACGGCGGG 1143
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      418 GlnSerTrpValAspLeuAlaAspLysTyrMetValGluAlaArgTrpPheTyrGlyGly 437

Qy     1144 CACAAACCAAGTTTGGAAGAGTATTTGGAGAACTCATGGCAGTCGATAAGTGGGCCCTGT 1203
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      438 HisLysProSerLeuGluGluTyrLeuGluAsnSerTrpGlnSerIleSerGlyProCys 457

Qy     1204 ATGTTAACGCACATATTCTTCCGAGTAACAGATTCGTTCACAAAGGAGACCGTCGACAGT 1263
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      458 MetLeuThrHisIlePhePheArgValThrAspSerPheThrLysGluThrValAspSer 477

Qy     1264 TTGTACAAATACCACGATTTAGTTCGTTGGTCATCCTTCGTTCTGCGGCTTGCTGATGAT 1323
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      478 LeuTyrLysTyrHisAspLeuValArgTrpSerSerPheValLeuArgLeuAlaAspAsp 497

Qy     1324 TTGGGAACCTCGGTGGAAGAGGTGAGCAGAGGGGATGTGCCGAAATCACTTCAGTGCTAC 1383
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      498 LeuGlyThrSerValGluGluValSerArgGlyAspValProLysSerLeuGlnCysTyr 517

Qy     1384 ATGAGTGACTACAATGCATCGGAGGCGGAGGCGCGGAAGCACGTGAAATGGCTGATAGCG 1443
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      518 MetSerAspTyrAsnAlaSerGluAlaGluAlaArgLysHisValLysTrpLeuIleAla 537

Qy     1444 GAGGTGTGGAAGAAGATGAATGCGGAGAGGGTGTCGAAGGATTCTCCATTCGGCAAAGAT 1503
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      538 GluValTrpLysLysMetAsnAlaGluArgValSerLysAspSerProPheGlyLysAsp 557

Qy     1504 TTTATAGGATGTGCAGTTGATTTAGGAAGGATGGCGCAGTTGATGTACCATAATGGAGAT 1563
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      558 PheIleGlyCysAlaValAspLeuGlyArgMetAlaGlnLeuMetTyrHisAsnGlyAsp 577

Qy     1564 GGGCACGGCACACAACACCCTATTATACATCAACAAATGACCAGAACCTTATTCGAGCCC 1623
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      578 GlyHisGlyThrGlnHisProIleIleHisGlnGlnMetThrArgThrLeuPheGluPro 597

Qy     1624 TTTGCA 1629
            ||||||
Db      598 PheAla 599

RESULT 4
US-09-900-797-22
; Sequence 22, Application US/09900797
; Publication No. US20030087406A1
; GENERAL INFORMATION:
;  APPLICANT: Chappell, Joseph
;  APPLICANT:  No. US20030087406A11, Joseph P.
;  APPLICANT:  Starks, Courtney M.
;  APPLICANT:  Manna, Kathleen R.
;  TITLE OF INVENTION: SYNTHASES
;  FILE REFERENCE: 07678-025001
;  CURRENT APPLICATION NUMBER: US/09/900,797
;  CURRENT FILING DATE:  2001-07-06
;  PRIOR APPLICATION NUMBER: US/09/398,395
;  PRIOR FILING DATE: 1999-09-17
```

```
;  PRIOR APPLICATION NUMBER: 60/130,628
;  PRIOR FILING DATE: 1999-04-22
;  PRIOR APPLICATION NUMBER: 60/150,262
;  PRIOR FILING DATE: 1999-08-23
;  NUMBER OF SEQ ID NOS: 58
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 22
;   LENGTH: 599
;   TYPE: PRT
;   ORGANISM: Mentha spicata
US-09-900-797-22

Alignment Scores:
Pred. No.:              4.17e-260     Length:        599
Score:                  2876.00       Matches:       542
Percent Similarity:     100.00%       Conservative:  0
Best Local Similarity:  100.00%       Mismatches:    0
Query Match:            98.90%        Indels:        0
DB:                     10            Gaps:          0

US-09-938-956-6 (1-1632) x US-09-900-797-22 (1-599)

Qy          4 AGACGATCCGGAAACTACAACCCTTCTCGTTGGGATGTCAACTTCATCCAATCGCTTCTC 63
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         58 ArgArgSerGlyAsnTyrAsnProSerArgTrpAspValAsnPheIleGlnSerLeuLeu 77

Qy         64 AGTGACTATAAGGAGGACAAACACGTGATTAGGGCTTCTGAGCTGGTCACTTTGGTGAAG 123
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         78 SerAspTyrLysGluAspLysHisValIleArgAlaSerGluLeuValThrLeuValLys 97

Qy        124 ATGGAACTGGAGAAAGAAACGGATCAAATTCGACAACTTGAGTTGATCGATGACTTGCAG 183
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         98 MetGluLeuGluLysGluThrAspGlnIleArgGlnLeuGluLeuIleAspAspLeuGln 117

Qy        184 AGGATGGGGCTGTCCGATCATTTCCAAAATGAGTTCAAAGAAATCTTGTCCTCTATATAT 243
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        118 ArgMetGlyLeuSerAspHisPheGlnAsnGluPheLysGluIleLeuSerSerIleTyr 137

Qy        244 CTCGACCATCACTATTACAAGAACCCTTTTCCAAAAGAAGAAAGGGATCTCTACTCCACA 303
              +|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        138 LeuAspHisHisTyrTyrLysAsnProPheProLysGluGluArgAspLeuTyrSerThr 157

Qy        304 TCTCTTGCATTTAGGCTCCTCAGAGAACATGGTTTTCAAGTCGCACAAGAGGTATTCGAT 363
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        158 SerLeuAlaPheArgLeuLeuArgGluHisGlyPheGlnValAlaGlnGluValPheAsp 177

Qy        364 AGTTTCAAGAACGAGGAGGGTGAGTTCAAAGAAAGCCTTAGCGACGACACCAGAGGATTG 423
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        178 SerPheLysAsnGluGluGlyGluPheLysGluSerLeuSerAspAspThrArgGlyLeu 197

Qy        424 TTGCAACTGTATGAAGCTTCCTTTCTGTTGACGGAAGGCGAAACCACGCTCGAGTCAGCG 483
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        198 LeuGlnLeuTyrGluAlaSerPheLeuLeuThrGluGlyGluThrThrLeuGluSerAla 217

Qy        484 AGGGAATTCGCCACCAAATTTTTGGAGGAAAAAGTGAACGAGGGTGGTGTTGATGGCGAC 543
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        218 ArgGluPheAlaThrLysPheLeuGluGluLysValAsnGluGlyGlyValAspGlyAsp 237

Qy        544 CTTTTAACAAGAATCGCATATTCTTTGGACATCCCTCTTCATTGGAGGATTAAAAGGCCA 603
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        238 LeuLeuThrArgIleAlaTyrSerLeuAspIleProLeuHisTrpArgIleLysArgPro 257

Qy        604 AATGCACCTGTGTGGATCGAATGGTATAGGAAGAGGCCCGACATGAATCCAGTAGTGTTG 663
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        258 AsnAlaProValTrpIleGluTrpTyrArgLysArgProAspMetAsnProValValLeu 277

Qy        664 GAGCTTGCCATACTCGACTTAAATATTGTTCAAGCACAATTTCAAGAAGAGCTCAAAGAA 723
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        278 GluLeuAlaIleLeuAspLeuAsnIleValGlnAlaGlnPheGlnGluGluLeuLysGlu 297

Qy        724 TCCTTCAGGTGGTGGAGAAATACTGGGTTTGTTGAGAAGCTGCCCTTCGCAAGGGATAGA 783
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        298 SerPheArgTrpTrpArgAsnThrGlyPheValGluLysLeuProPheAlaArgAspArg 317

Qy        784 CTGGTGGAATGCTACTTTTGGAATACTGGGATCATCGAGCCACGTCAGCATGCAAGTGCA 843
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        318 LeuValGluCysTyrPheTrpAsnThrGlyIleIleGluProArgGlnHisAlaSerAla 337

Qy        844 AGGATAATGATGGGCAAAGTCAACGCTCTGATTACGGTGATCGATGATATTTATGATGTC 903
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        338 ArgIleMetMetGlyLysValAsnAlaLeuIleThrValIleAspAspIleTyrAspVal 357

Qy        904 TATGGCACCTTAGAAGAACTCGAACAATTCACTGACCTCATTCGAAGATGGGATATAAAC 963
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        358 TyrGlyThrLeuGluGluLeuGluGlnPheThrAspLeuIleArgArgTrpAspIleAsn 377

Qy        964 TCAATCGACCAACTTCCCGATTACATGCAACTGTGCTTTCTTGCACTCAACAACTTCGTC 1023
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        378 SerIleAspGlnLeuProAspTyrMetGlnLeuCysPheLeuAlaLeuAsnAsnPheVal 397

Qy       1024 GATGATACATCGTACGATGTTATGAAGGAGAAAGGCGTCAACGTTATACCCTACCTGCGG 1083
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        398 AspAspThrSerTyrAspValMetLysGluLysGlyValAsnValIleProTyrLeuArg 417

Qy       1084 CAATCGTGGGTTGATTTGGCGGATAAGTATATGGTAGAGGCACGGTGGTTCTACGGCGGG 1143
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        418 GlnSerTrpValAspLeuAlaAspLysTyrMetValGluAlaArgTrpPheTyrGlyGly 437

Qy       1144 CACAAACCAAGTTTGGAAGAGTATTTGGAGAACTCATGGCAGTCGATAAGTGGCCCCTGT 1203
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        438 HisLysProSerLeuGluGluTyrLeuGluAsnSerTrpGlnSerIleSerGlyProCys 457

Qy       1204 ATGTTAACGCACATATTCTTCCGAGTAACAGATTCGTTCACAAAGGAGACCGTCGACAGT 1263
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        458 MetLeuThrHisIlePhePheArgValThrAspSerPheThrLysGluThrValAspSer 477

Qy       1264 TTGTACAAATACCACGATTTAGTTCGTTGGTCATCCTTCGTTCTGCGGCTTGCTGATGAT 1323
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        478 LeuTyrLysTyrHisAspLeuValArgTrpSerSerPheValLeuArgLeuAlaAspAsp 497

Qy       1324 TTGGGAACCTCGGTGGAAGAGGTGAGCAGAGGGGATGTGCCGAAATCACTTCAGTGCTAC 1383
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        498 LeuGlyThrSerValGluGluValSerArgGlyAspValProLysSerLeuGlnCysTyr 517

Qy       1384 ATGAGTGACTACAATGCATCGGAGGCGGAGGCGCGGAAGCACGTGAAATGGCTGATAGCG 1443
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        518 MetSerAspTyrAsnAlaSerGluAlaGluAlaArgLysHisValLysTrpLeuIleAla 537

Qy       1444 GAGGTGTGGAAGAAGATGAATGCGGAGAGGGTGTCGAAGGATTCTCCATTCGGCAAAGAT 1503
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        538 GluValTrpLysLysMetAsnAlaGluArgValSerLysAspSerProPheGlyLysAsp 557

Qy       1504 TTTATAGGATGTGCAGTTGATTTAGGAAGGATGGCGCAGTTGATGTACCATAATGGAGAT 1563
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        558 PheIleGlyCysAlaValAspLeuGlyArgMetAlaGlnLeuMetTyrHisAsnGlyAsp 577

Qy       1564 GGGCACGGCACACAACACCCTATTATACATCAACAAATGACCAGAACCTTATTCGAGCCC 1623
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        578 GlyHisGlyThrGlnHisProIleIleHisGlnGlnMetThrArgThrLeuPheGluPro 597

Qy       1624 TTTGCA 1629
              ||||||
Db        598 PheAla 599

RESULT 5
US-09-887-586A-24
; Sequence 24, Application US/09887586A
; Patent No. US20020094556A1
; GENERAL INFORMATION:
;  APPLICANT: Chappell, Joseph
;  APPLICANT:  No. US20020094556A11, Joseph P.
;  APPLICANT:  Starks, Courtney M.
;  APPLICANT:  Manna, Kathleen R.
;  TITLE OF INVENTION: SYNTHASES
;  FILE REFERENCE: 07678-025001
;  CURRENT APPLICATION NUMBER: US/09/887,586A
```

```
;  CURRENT FILING DATE:  2001-06-22
;  PRIOR APPLICATION NUMBER: 09/398,395
;  PRIOR FILING DATE: 1999-09-17
;  PRIOR APPLICATION NUMBER: 60/130,628
;  PRIOR FILING DATE: 1999-04-22
;  PRIOR APPLICATION NUMBER: 60/150,262
;  PRIOR FILING DATE: 1999-08-23
;  NUMBER OF SEQ ID NOS: 58
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 24
;   LENGTH: 591
;   TYPE: PRT
;   ORGANISM: Salvia officinalis
US-09-887-586A-24

Alignment Scores:
Pred. No.:              9.03e-144     Length:        591
Score:                  1633.50       Matches:       305
Percent Similarity:     75.96%        Conservative:  109
Best Local Similarity:  55.96%        Mismatches:    116
Query Match:            56.17%        Indels:        15
DB:                     9             Gaps:          7

US-09-938-956-6 (1-1632) x US-09-887-586A-24 (1-591)

Qy      4 AGACGATCCGGAAACTACAACCCTTCTCGTTGGGATGTCAACTTCATCCAATCGCTTCTC 63
Db     57 ArgArgThrGlyGlyTyrGlnProThrLeuTrpAspPheSerThrIleGlnLeuPheAsp 76

Qy     64 AGTGACTATAAGGAGGACAAACACGTGATTAGGGCTTCTGAGCTGGTCACTTTGGTGAAG 123
Db     77 SerGluTyrLysGluGluLysHisLeuMetArgAlaAlaGlyMetIleAlaGlnValAsn 96

Qy    124 ATGGAACTGGAGAAAGAAACGGATCAAATTCGACAACTTGAGTTGATCGATGACTTGCAG 183
Db     97 MetLeuLeuGlnGluGluValAspSerIleGlnArgLeuGluLeuIleAspAspLeuArg 116

Qy    184 AGGATGGGGCTGTCCGATCATTTCCAAAATGAGTTCAAAGAAATCTTGTCCTCTATATAT 243
Db    117 ArgLeuGlyIleSerCysHisPheAspArgGluIleValGluIleLeuAsnSerLysTyr 136

Qy    244 CTCGACCATCACTATTACAAGAACCCTTTTCCAAAAGAAGAAAGGGATCTCTACTCCACA 303
Db    137 ------------TyrThrAsnAsnGluIle------AspGluSerAspLeuTyrSerThr 150

Qy    304 TCTCTTGCATTTAGGCTCCTCAGAGAACATGGTTTTCAAGTCGCACAAGAGGTATTCGAT 363
Db    151 AlaLeuArgPheLysLeuLeuArgGlnTyrAspPheSerValSerGlnGluValPheAsp 170

Qy    364 AGTTTCAAGAACGAGGAGGGT---GAGTTCAAAGAAAGCCTTAGCGACGACACCAGAGGA 420
Db    171 CysPheLysAsnAspLysGlyThrAspPheLysProSerLeuValAspAspThrArgGly 190

Qy    421 TTGTTGCAACTGTATGAAGCTTCCTTTCTGTTGACGGAAGGCGAAACCACGCTCGAGTCA 480
Db    191 LeuLeuGlnLeuTyrGluAlaSerPheLeuSerAlaGlnGlyGluGluThrLeuHisLeu 210

Qy    481 GCGAGGGAATTCGCCACCAAATTTTTGGAGGAAAAAGTGAACGAGGGTGGTGTTGATGGC 540
Db    211 AlaArgAspPheAlaThrLysPheLeuHisLysArgValLeu---------ValAspLys 227

Qy    541 GAC------CTTTTAACAAGAATCGCATATTCTTTGGACATCCCTCTTCATTGGAGGATT 594
Db    228 AspIleAsnLeuLeuSerSerIleGluArgAlaLeuGluLeuProThrHisTrpArgVal 247

Qy    595 AAAAGGCCAAATGCACCTGTGTGGATCGAATGGTATAGGAAGAGGCCCGACATGAATCCA 654
Db    248 GlnMetProAsnAlaArgSerPheIleAspAlaTyrLysArgArgProAspMetAsnPro 267

Qy    655 GTAGTGTTGGAGCTTGCCATACTCGACTTAAATATTGTTCAAGCACAATTTCAAGAAGAG 714
Db    268 ThrValLeuGluLeuAlaLysLeuAspPheAsnMetValGlnAlaGlnPheGlnGlnGlu 287

Qy    715 CTCAAAGAATCCTTCAGGTGGTGGAGAAATACTGGGTTTGTTGAGAAGCTGCCCTTCGCA 774
Db    288 LeuLysGluAlaSerArgTrpTrpAsnSerThrGlyLeuValHisGluLeuProPheVal 307

Qy    775 AGGGATAGACTGGTGGAATGCTACTTTTGGAATACTGGGATCATCGAGCCACGTCAGCAT 834
Db    308 ArgAspArgIleValGluCysTyrTyrTrpThrThrGlyValValGluArgArgGluHis 327

Qy    835 GCAAGTGCAAGGATAATGATGGGCAAAGTCAACGCTCTGATTACGGTGATCGATGATATT 894
Db    328 GlyTyrGluArgIleMetLeuThrLysIleAsnAlaLeuValThrThrIleAspAspVal 347

Qy    895 TATGATGTCTATGGCACCTTAGAAGAACTCGAACAATTCACTGACCTCATTCGAAGATGG 954
Db    348 PheAspIleTyrGlyThrLeuGluGluLeuGlnLeuPheThrThrAlaIleGlnArgTrp 367

Qy    955 GATATAAACTCAATCGACCAACTTCCCGATTACATGCAACTGTGCTTTCTTGCACTCAAC 1014
Db    368 AspIleGluSerMetLysGlnLeuProProTyrMetGlnIleCysTyrLeuAlaLeuPhe 387

Qy   1015 AACTTCGTCGATGATACATCGTACGATGTTATGAAGGAGAAAGGCGTCAACGTTATACCC 1074
Db    388 AsnPheValAsnGluMetAlaTyrAspThrLeuArgAspLysGlyPheAsnSerThrPro 407

Qy   1075 TACCTGCGGCAATCGTGGGTTGATTTGGCGGATAAGTATATGGTAGAGGCACGGTGGTTC 1134
Db    408 TyrLeuArgLysAlaTrpValAspLeuValGluSerTyrLeuIleGluAlaLysTrpTyr 427

Qy   1135 TACGGCGGGCACAAACCAAGTTTGGAAGAGTATTTGGAGAACTCATGGCAGTCGATAAGT 1194
Db    428 TyrMetGlyHisLysProSerLeuGluGluTyrMetLysAsnSerTrpIleSerIleGly 447

Qy   1195 GGGCCCTGTATGTTAACGCACATATTCTTCCGAGTAACAGATTCGTTCACAAAGGAGACC 1254
Db    448 GlyIleProIleLeuSerHisLeuPhePheArgLeuThrAspSerIleGluGluGluAsp 467

Qy   1255 GTCGACAGTTTGTACAAATACCACGATTTAGTTCGTTGGTCATCCTTCGTTCTGCGGCTT 1314
Db    468 AlaGluSerMetHisLysTyrHisAspIleValArgAlaSerCysThrIleLeuArgLeu 487

Qy   1315 GCTGATGATTTGGGAACCTCGGTGGAAGAGGTGAGCAGAGGGGATGTGCCGAAATCACTT 1374
Db    488 AlaAspAspMetGlyThrSerLeuAspGluValGluArgGlyAspValProLysSerVal 507

Qy   1375 CAGTGCTACATGAGTGACTACAATGCATCGGAGGCGGAGGCGCGGAAGCACGTGAAATGG 1434
Db    508 GlnCysTyrMetAsnGluLysAsnAlaSerGluGluGluAlaArgGluHisValArgSer 527

Qy   1435 CTGATAGCGGAGGTGTGGAAGAAGATGAATGCGGAGAGGGTGTCGAAGGATTCTCCATTC 1494
Db    528 LeuIleAspGlnThrTrpLysMetMetAsnLysGluMetMetThr------SerSerPhe 545

Qy   1495 GGCAAAGATTTTATAGGATGTGCAGTTGATTTAGGAAGGATGGCGCAGTTGATGTAC--- 1551
Db    546 SerLysTyrPheValGlnValSerAlaAsnLeuAlaArgMetAlaGlnTrpIleTyrGln 565

Qy   1552 CATAATGGAGATGGGCACGGCACACAACACCCTATTATACATCAACAAATGACCAGAACC 1611
Db    566 HisGluSerAspGlyPheGlyMetGlnHisSerLeuValAsnLysMetLeuArgGlyLeu 585

Qy   1612 TTATTCGAGCCCTTT 1626
Db    586 LeuPheAspArgTyr 590

RESULT 6
US-09-903-012-24
; Sequence 24, Application US/09903012
; Patent No. US20020094557A1
; GENERAL INFORMATION:
;  APPLICANT: Chappell, Joseph
;  APPLICANT:  No. US20020094557A11, Joseph P.
;  APPLICANT:  Starks, Courtney M.
;  APPLICANT:  Manna, Kathleen R.
```

```
;  TITLE OF INVENTION: SYNTHASES
;  FILE REFERENCE: 07678-025001
;  CURRENT APPLICATION NUMBER: US/09/903,012
;  CURRENT FILING DATE:  2001-07-11
;  PRIOR APPLICATION NUMBER: 09/398,395
;  PRIOR FILING DATE: 1999-09-17
;  PRIOR APPLICATION NUMBER: 60/100,993
;  PRIOR FILING DATE: 1998-09-18
;  PRIOR APPLICATION NUMBER: 60/130,628
;  PRIOR FILING DATE: 1999-04-22
;  PRIOR APPLICATION NUMBER: 60/150,262
;  PRIOR FILING DATE: 1999-08-23
;  NUMBER OF SEQ ID NOS: 58
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 24
;   LENGTH: 591
;   TYPE: PRT
;   ORGANISM: Salvia officinalis
US-09-903-012-24

Alignment Scores:
Pred. No.:              9.03e-144     Length:         591
Score:                  1633.50       Matches:        305
Percent Similarity:     75.96%        Conservative:   109
Best Local Similarity:  55.96%        Mismatches:     116
Query Match:            56.17%        Indels:         15
DB:                     9             Gaps:           7

US-09-938-956-6 (1-1632) x US-09-903-012-24 (1-591)

Qy        4 AGACGATCCGGAAACTACAACCCTTCTCGTTGGGATGTCAACTTCATCCAATCGCTTCTC 63
Db       57 ArgArgThrGlyGlyTyrGlnProThrLeuTrpAspPheSerThrIleGlnLeuPheAsp 76

Qy       64 AGTGACTATAAGGAGGACAAACACGTGATTAGGGCTTCTGAGCTGGTCACTTTGGTGAAG 123
Db       77 SerGluTyrLysGluGluLysHisLeuMetArgAlaAlaGlyMetIleAlaGlnValAsn 96

Qy      124 ATGGAACTGGAGAAAGAAACGGATCAAATTCGACAACTTGAGTTGATCGATGACTTGCAG 183
Db       97 MetLeuLeuGlnGluGluValAspSerIleGlnArgLeuGluLeuIleAspAspLeuArg 116

Qy      184 AGGATGGGGCTGTCCGATCATTTCCAAAATGAGTTCAAAGAAATCTTGTCCTCTATATAT 243
Db      117 ArgLeuGlyIleSerCysHisPheAspArgGluIleValGluIleLeuAsnSerLysTyr 136

Qy      244 CTCGACCATCACTATTACAAGAACCCTTTTCCAAAAGAAGAAAGGGATCTCTACTCCACA 303
Db      137 ------------TyrThrAsnAsnGluIle------AspGluSerAspLeuTyrSerThr 150

Qy      304 TCTCTTGCATTTAGGCTCCTCAGAGAACATGGTTTTCAAGTCGCACAAGAGGTATTCGAT 363
Db      151 AlaLeuArgPheLysLeuLeuArgGlnTyrAspPheSerValSerGlnGluValPheAsp 170

Qy      364 AGTTTCAAGAACGAGGAGGGT---GAGTTCAAAGAAAGCCTTAGCGACGACACCAGAGGA 420
Db      171 CysPheLysAsnAspLysGlyThrAspPheLysProSerLeuValAspAspThrArgGly 190

Qy      421 TTGTTGCAACTGTATGAAGCTTCCTTTCTGTTGACGGAAGGCGAAACCACGCTCGAGTCA 480
Db      191 LeuLeuGlnLeuTyrGluAlaSerPheLeuSerAlaGlnGlyGluGluThrLeuHisLeu 210

Qy      481 GCGAGGGAATTCGCCACCAAATTTTTGGAGGAAAAAGTGAACGAGGGTGGTGTTGATGGC 540
Db      211 AlaArgAspPheAlaThrLysPheLeuHisLysArgValLeu---------ValAspLys 227

Qy      541 GAC------CTTTTAACAAGAATCGCATATTCTTTGGACATCCCTCTTCATTGGAGGATT 594
Db      228 AspIleAsnLeuLeuSerSerIleGluArgAlaLeuGluLeuProThrHisTrpArgVal 247

Qy      595 AAAAGGCCAAATGCACCTGTGTGGATCGAATGGTATAGGAAGAGGCCCGACATGAATCCA 654
Db      248 GlnMetProAsnAlaArgSerPheIleAspAlaTyrLysArgArgProAspMetAsnPro 267

Qy      655 GTAGTGTTGGAGCTTGCCATACTCGACTTAAATATTGTTCAAGCACAATTTCAAGAAGAG 714
Db      268 ThrValLeuGluLeuAlaLysLeuAspPheAsnMetValGlnAlaGlnPheGlnGlnGlu 287

Qy      715 CTCAAAGAATCCTTCAGGTGGTGGAGAAATACTGGGTTTGTTGAGAAGCTGCCCTTCGCA 774
Db      288 LeuLysGluAlaSerArgTrpTrpAsnSerThrGlyLeuValHisGluLeuProPheVal 307

Qy      775 AGGGATAGACTGGTGGAATGCTACTTTTGGAATACTGGGATCATCGAGCCACGTCAGCAT 834
Db      308 ArgAspArgIleValGluCysTyrTyrTrpThrThrGlyValValGluArgArgGluHis 327

Qy      835 GCAAGTGCAAGGATAATGATGGGCAAAGTCAACGCTCTGATTACGGTGATCGATGATATT 894
Db      328 GlyTyrGluArgIleMetLeuThrLysIleAsnAlaLeuValThrThrIleAspAspVal 347

Qy      895 TATGATGTCTATGGCACCTTAGAAGAACTCGAACAATTCACTGACCTCATTCGAAGATGG 954
Db      348 PheAspIleTyrGlyThrLeuGluGluLeuGlnLeuPheThrThrAlaIleGlnArgTrp 367

Qy      955 GATATAAACTCAATCGACCAACTTCCCGATTACATGCAACTGTGCTTTCTTGCACTCAAC 1014
Db      368 AspIleGluSerMetLysGlnLeuProProTyrMetGlnIleCysTyrLeuAlaLeuPhe 387

Qy     1015 AACTTCGTCGATGATACATCGTACGATGTTATGAAGGAGAAAGGCGTCAACGTTATACCC 1074
Db      388 AsnPheValAsnGluMetAlaTyrAspThrLeuArgAspLysGlyPheAsnSerThrPro 407

Qy     1075 TACCTGCGGCAATCGTGGGTTGATTTGGCGGATAAGTATATGGTAGAGGCACGGTGGTTC 1134
Db      408 TyrLeuArgLysAlaTrpValAspLeuValGluSerTyrLeuIleGluAlaLysTrpTyr 427

Qy     1135 TACGGCGGGCACAAACCAAGTTTGGAAGAGTATTTGGAGAACTCATGGCAGTCGATAAGT 1194
Db      428 TyrMetGlyHisLysProSerLeuGluGluTyrMetLysAsnSerTrpIleSerIleGly 447

Qy     1195 GGGCCCTGTATGTTAACGCACATATTCTTCCGAGTAACAGATTCGTTCACAAAGGAGACC 1254
Db      448 GlyIleProIleLeuSerHisLeuPhePheArgLeuThrAspSerIleGluGluGluAsp 467

Qy     1255 GTCGACAGTTTGTACAAATACCACGATTTAGTTCGTTGGTCATCCTTCGTTCTGCGGCTT 1314
Db      468 AlaGluSerMetHisLysTyrHisAspIleValArgAlaSerCysThrIleLeuArgLeu 487

Qy     1315 GCTGATGATTTGGGAACCTCGGTGGAAGAGGTGAGCAGAGGGGATGTGCCGAAATCACTT 1374
Db      488 AlaAspAspMetGlyThrSerLeuAspGluValGluArgGlyAspValProLysSerVal 507

Qy     1375 CAGTGCTACATGAGTGACTACAATGCATCGGAGGCGGAGGCGCGGAAGCACGTGAAATGG 1434
Db      508 GlnCysTyrMetAsnGluLysAsnAlaSerGluGluGluAlaArgGluHisValArgSer 527

Qy     1435 CTGATAGCGGAGGTGTGGAAGAAGATGAATGCGGAGAGGGTGTCGAAGGATTCTCCATTC 1494
Db      528 LeuIleAspGlnThrTrpLysMetMetAsnLysGluMetMetThr------SerSerPhe 545

Qy     1495 GGCAAAGATTTTATAGGATGTGCAGTTGATTTAGGAAGGATGGCGCAGTTGATGTAC--- 1551
Db      546 SerLysTyrPheValGlnValSerAlaAsnLeuAlaArgMetAlaGlnTrpIleTyrGln 565

Qy     1552 CATAATGGAGATGGGCACGGCACACAACACCCTATTATACATCAACAAATGACCAGAACC 1611
Db      566 HisGluSerAspGlyPheGlyMetGlnHisSerLeuValAsnLysMetLeuArgGlyLeu 585

Qy     1612 TTATTCGAGCCCTTT 1626
Db      586 LeuPheAspArgTyr 590

RESULT 7
US-09-900-797-24
; Sequence 24, Application US/09900797
; Publication No. US20030087406A1
```

```
; GENERAL INFORMATION:
;  APPLICANT: Chappell, Joseph
;  APPLICANT:  No. US20030087406A11, Joseph P.
;  APPLICANT:  Starks, Courtney M.
;  APPLICANT:  Manna, Kathleen R.
;  TITLE OF INVENTION: SYNTHASES
;  FILE REFERENCE: 07678-025001
;  CURRENT APPLICATION NUMBER: US/09/900,797
;  CURRENT FILING DATE:  2001-07-06
;  PRIOR APPLICATION NUMBER: US/09/398,395
;  PRIOR FILING DATE: 1999-09-17
;  PRIOR APPLICATION NUMBER: 60/130,628
;  PRIOR FILING DATE: 1999-04-22
;  PRIOR APPLICATION NUMBER: 60/150,262
;  PRIOR FILING DATE: 1999-08-23
;  NUMBER OF SEQ ID NOS: 58
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 24
;   LENGTH: 591
;   TYPE: PRT
;   ORGANISM: Salvia officinalis
US-09-900-797-24

Alignment Scores:
Pred. No.:              9.03e-144     Length:          591
Score:                  1633.50       Matches:         305
Percent Similarity:     75.96%        Conservative:    109
Best Local Similarity:  55.96%        Mismatches:      116
Query Match:            56.17%        Indels:          15
DB:                     10            Gaps:            7

US-09-938-956-6 (1-1632) x US-09-900-797-24 (1-591)

Qy          4 AGACGATCCGGAAACTACAACCCTTCTCGTTGGGATGTCAACTTCATCCAATCGCTTCTC 63
Db         57 ArgArgThrGlyGlyTyrGlnProThrLeuTrpAspPheSerThrIleGlnLeuPheAsp 76

Qy         64 AGTGACTATAAGGAGGACAAACACGTGATTAGGGCTTCTGAGCTGGTCACTTTGGTGAAG 123
Db         77 SerGluTyrLysGluGluLysHisLeuMetArgAlaAlaGlyMetIleAlaGlnValAsn 96

Qy        124 ATGGAACTGGAGAAAGAAACGGATCAAATTCGACAACTTGAGTTGATCGATGACTTGCAG 183
Db         97 MetLeuLeuGlnGluGluValAspSerIleGlnArgLeuGluLeuIleAspAspLeuArg 116

Qy        184 AGGATGGGGCTGTCCGATCATTTCCAAAATGAGTTCAAAGAAATCTTGTCCTCTATATAT 243
Db        117 ArgLeuGlyIleSerCysHisPheAspArgGluIleValGluIleLeuAsnSerLysTyr 136

Qy        244 CTCGACCATCACTATTACAAGAACCCTTTTCCAAAAGAAGAAAGGGATCTCTACTCCACA 303
Db        137 ------------TyrThrAsnAsnGluIle------AspGluSerAspLeuTyrSerThr 150

Qy        304 TCTCTTGCATTTAGGCTCCTCAGAGAACATGGTTTTCAAGTCGCACAAGAGGTATTCGAT 363
Db        151 AlaLeuArgPheLysLeuLeuArgGlnTyrAspPheSerValSerGlnGluValPheAsp 170

Qy        364 AGTTTCAAGAACGAGGAGGGT---GAGTTCAAAGAAAGCCTTAGCGACGACACCAGAGGA 420
Db        171 CysPheLysAsnAspLysGlyThrAspPheLysProSerLeuValAspAspThrArgGly 190

Qy        421 TTGTTGCAACTGTATGAAGCTTCCTTTCTGTTGACGGAAGGCGAAACCACGCTCGAGTCA 480
Db        191 LeuLeuGlnLeuTyrGluAlaSerPheLeuSerAlaGlnGlyGluGluThrLeuHisLeu 210

Qy        481 GCGAGGGAATTCGCCACCAAATTTTTGGAGGAAAAAGTGAACGAGGGTGGTGTTGATGGC 540
Db        211 AlaArgAspPheAlaThrLysPheLeuHisLysArgValLeu---------ValAspLys 227

Qy        541 GAC------CTTTTAACAAGAATCGCATATTCTTTGGACATCCCTCTTCATTGGAGGATT 594
Db        228 AspIleAsnLeuLeuSerSerIleGluArgAlaLeuGluLeuProThrHisTrpArgVal 247

Qy        595 AAAAGGCCAAATGCACCTGTGTGGATCGAATGGTATAGGAAGAGGCCCGACATGAATCCA 654
Db        248 GlnMetProAsnAlaArgSerPheIleAspAlaTyrLysArgArgProAspMetAsnPro 267

Qy        655 GTAGTGTTGGAGCTTGCCATACTCGACTTAAATATTGTTCAAGCACAATTTCAAGAAGAG 714
Db        268 ThrValLeuGluLeuAlaLysLeuAspPheAsnMetValGlnAlaGlnPheGlnGlnGlu 287

Qy        715 CTCAAAGAATCCTTCAGGTGGTGGAGAAATACTGGGTTTGTTGAGAAGCTGCCCTTCGCA 774
Db        288 LeuLysGluAlaSerArgTrpTrpAsnSerThrGlyLeuValHisGluLeuProPheVal 307

Qy        775 AGGGATAGACTGGTGGAATGCTACTTTTGGAATACTGGGATCATCGAGCCACGTCAGCAT 834
Db        308 ArgAspArgIleValGluCysTyrTyrTrpThrThrGlyValValGluArgArgGluHis 327

Qy        835 GCAAGTGCAAGGATAATGATGGGCAAAGTCAACGCTCTGATTACGGTGATCGATGATATT 894
Db        328 GlyTyrGluArgIleMetLeuThrLysIleAsnAlaLeuValThrThrIleAspAspVal 347

Qy        895 TATGATGTCTATGGCACCTTAGAAGAACTCGAACAATTCACTGACCTCATTCGAAGATGG 954
Db        348 PheAspIleTyrGlyThrLeuGluGluLeuGlnLeuPheThrThrAlaIleGlnArgTrp 367

Qy        955 GATATAAACTCAATCGACCAACTTCCCGATTACATGCAACTGTGCTTTCTTGCACTCAAC 1014
Db        368 AspIleGluSerMetLysGlnLeuProProTyrMetGlnIleCysTyrLeuAlaLeuPhe 387

Qy       1015 AACTTCGTCGATGATACATCGTACGATGTTATGAAGGAGAAAGGCGTCAACGTTATACCC 1074
Db        388 AsnPheValAsnGluMetAlaTyrAspThrLeuArgAspLysGlyPheAsnSerThrPro 407

Qy       1075 TACCTGCGGCAATCGTGGGTTGATTTGGCGGATAAGTATATGGTAGAGGCACGGTGGTTC 1134
Db        408 TyrLeuArgLysAlaTrpValAspLeuValGluSerTyrLeuIleGluAlaLysTrpTyr 427

Qy       1135 TACGGCGGGCACAAACCAAGTTTGGAAGAGTATTTGGAGAACTCATGGCAGTCGATAAGT 1194
Db        428 TyrMetGlyHisLysProSerLeuGluGluTyrMetLysAsnSerTrpIleSerIleGly 447

Qy       1195 GGGCCCTGTATGTTAACGCACATATTCTTCCGAGTAACAGATTCGTTCACAAAGGAGACC 1254
Db        448 GlyIleProIleLeuSerHisLeuPhePheArgLeuThrAspSerIleGluGluGluAsp 467

Qy       1255 GTCGACAGTTTGTACAAATACCACGATTTAGTTCGTTGGTCATCCTTCGTTCTGCGGCTT 1314
Db        468 AlaGluSerMetHisLysTyrHisAspIleValArgAlaSerCysThrIleLeuArgLeu 487

Qy       1315 GCTGATGATTTGGGAACCTCGGTGGAAGAGGTGAGCAGAGGGGATGTGCCGAAATCACTT 1374
Db        488 AlaAspAspMetGlyThrSerLeuAspGluValGluArgGlyAspValProLysSerVal 507

Qy       1375 CAGTGCTACATGAGTGACTACAATGCATCGGAGGCGGAGGCGCGGAAGCACGTGAAATGG 1434
Db        508 GlnCysTyrMetAsnGluLysAsnAlaSerGluGluGluAlaArgGluHisValArgSer 527

Qy       1435 CTGATAGCGGAGGTGTGGAAGAAGATGAATGCGGAGAGGGTGTCGAAGGATTCTCCATTC 1494
Db        528 LeuIleAspGlnThrTrpLysMetMetAsnLysGluMetMetThr------SerSerPhe 545

Qy       1495 GGCAAAGATTTTATAGGATGTGCAGTTGATTTAGGAAGGATGGCGCAGTTGATGTAC--- 1551
Db        546 SerLysTyrPheValGlnValSerAlaAsnLeuAlaArgMetAlaGlnTrpIleTyrGln 565

Qy       1552 CATAATGGAGATGGGCACGGCACACAACACCCTATTATACATCAACAAATGACCAGAACC 1611
Db        566 HisGluSerAspGlyPheGlyMetGlnHisSerLeuValAsnLysMetLeuArgGlyLeu 585

Qy       1612 TTATTCGAGCCCTTT 1626
Db        586 LeuPheAspArgTyr 590

RESULT 8
```

```
US-09-887-586A-54
; Sequence 54, Application US/09887586A
; Patent No. US20020094556A1
; GENERAL INFORMATION:
;  APPLICANT: Chappell, Joseph
;  APPLICANT:  No. US20020094556A11, Joseph P.
;  APPLICANT:  Starks, Courtney M.
;  APPLICANT:  Manna, Kathleen R.
;  TITLE OF INVENTION: SYNTHASES
;  FILE REFERENCE: 07678-025001
;  CURRENT APPLICATION NUMBER: US/09/887,586A
;  CURRENT FILING DATE:  2001-06-22
;  PRIOR APPLICATION NUMBER: 09/398,395
;  PRIOR FILING DATE: 1999-09-17
;  PRIOR APPLICATION NUMBER: 60/130,628
;  PRIOR FILING DATE: 1999-04-22
;  PRIOR APPLICATION NUMBER: 60/150,262
;  PRIOR FILING DATE: 1999-08-23
;  NUMBER OF SEQ ID NOS: 58
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 54
;   LENGTH: 590
;   TYPE: PRT
;   ORGANISM: Salvia officinalis
US-09-887-586A-54

Alignment Scores:
Pred. No.:               2.79e-136     Length:         590
Score:                   1553.50       Matches:        291
Percent Similarity:      72.48%        Conservative:   104
Best Local Similarity:   53.39%        Mismatches:     141
Query Match:             53.42%        Indels:         9
DB:                      9             Gaps:           3

US-09-938-956-6 (1-1632) x US-09-887-586A-54 (1-590)

Qy      1 ATGAGACGATCCGGAAACTACAACCCTTCTCGTTGGGATGTCAACTTCATCCAATCGCTT 60
Db     51 IleArgArgSerGlyAspTyrGlnProSerLeuTrpAspPheAsnTyrIleGlnSerLeu 70

Qy     61 CTCAGTGACTATAAGGAGGACAAACACGTGATTAGGGCTTCTGAGCTGGTCACTTTGGTG 120
Db     71 AsnThrProTyrLysGluGlnArgHisPheAsnArgGlnAlaGluLeuIleMetGlnVal 90

Qy    121 AAGATGGAACTGGAGAAAGAAACGGATCAAATTCGACAACTTGAGTTGATCGATGACTTG 180
Db     91 ArgMetLeuLeuLysValLysMetGluAlaIleGlnGlnLeuGluLeuIleAspAspLeu 110

Qy    181 CAGAGGATGGGGCTGTCCGATCATTTCCAAAATGAGTTCAAAGAAATCTTGTCCTCTATA 240
Db    111 GlnTyrLeuGlyLeuSerTyrPhePheGlnAspGluIleLysGlnIleLeuSerSerIle 130

Qy    241 TATCTCGACCATCACTATTACAAGAACCCTTTTCCAAAAGAAGAAAGGGATCTCTACTCC 300
Db    131 HisAsnGluProArgTyrPheHisAsn------------------AsnAspLeuTyrPhe 144

Qy    301 ACATCTCTTGCATTTAGGCTCCTCAGAGAACATGGTTTTCAAGTCGCACAAGAGGTATTC 360
Db    145 ThrAlaLeuGlyPheArgIleLeuArgGlnHisGlyPheAsnValSerGluAspValPhe 164

Qy    361 GATAGTTTCAAGAACGAGGAG---GGTGAGTTCAAAGAAAGCCTTAGCGACGACACCAGA 417
Db    165 AspCysPheLysIleGluLysCysSerAspPheAsnAlaAsnLeuAlaGlnAspThrLys 184

Qy    418 GGATTGTTGCAACTGTATGAAGCTTCCTTTCTGTTGACGGAAGGCGAAACCACGCTCGAG 477
Db    185 GlyMetLeuGlnLeuTyrGluAlaSerPheLeuLeuArgGluGlyGluAspThrLeuGlu 204

Qy    478 TCAGCGAGGGAATTCGCCACCAAATTTTTGGAGGAAAAAGTGAACGAGGGTGGT------ 531
Db    205 LeuAlaArgArgPheSerThrArgSerLeuArgGluLysPheAspGluGlyGlyAspGlu 224

Qy    532 GTTGATGGCGACCTTTTAACAAGAATCGCATATTCTTTGGACATCCCTCTTCATTGGAGG 591
Db    225 IleAspGluAspLeuSerSerTrpIleArgHisSerLeuAspLeuProLeuHisTrpArg 244

Qy    592 ATTAAAAGGCCAAATGCACCTGTGTGGATCGAATGGTATAGGAAGAGGCCCGACATGAAT 651
Db    245 ValGlnGlyLeuGluAlaArgTrpPheLeuAspAlaTyrAlaArgArgProAspMetAsn 264

Qy    652 CCAGTAGTGTTGGAGCTTGCCATACTCGACTTAAATATTGTTCAAGCACAATTTCAAGAA 711
Db    265 ProLeuIlePheLysLeuAlaLysLeuAsnPheAsnIleValGlnAlaThrTyrGlnGlu 284

Qy    712 GAGCTCAAAGAATCCTTCAGGTGGTGGAGAAATACTGGGTTTGTTGAGAAGCTGCCCTTC 771
Db    285 GluLeuLysAspIleSerArgTrpTrpAsnSerSerCysLeuAlaGluLysLeuProPhe 304

Qy    772 GCAAGGGATAGACTGGTGGAATGCTACTTTTGGAATACTGGGATCATCGAGCCACGTCAG 831
Db    305 ValArgAspArgIleValGluCysPhePheTrpAlaIleAlaAlaPheGluProHisGln 324

Qy    832 CATGCAAGTGCAAGGATAATGATGGGCAAAGTCAACGCTCTGATTACGGTGATCGATGAT 891
Db    325 TyrSerTyrGlnArgLysMetAlaAlaValIleIleThrPheIleThrIleIleAspAsp 344

Qy    892 ATTTATGATGTCTATGGCACCTTAGAAGAACTCGAACAATTCACTGACCTCATTCGAAGA 951
Db    345 ValTyrAspValTyrGlyThrIleGluGluLeuGluLeuLeuThrAspMetIleArgArg 364

Qy    952 TGGGATATAAACTCAATCGACCAACTTCCCGATTACATGCAACTGTGCTTTCTTGCACTC 1011
Db    365 TrpAspAsnLysSerIleSerGlnLeuProTyrTyrMetGlnValCysTyrLeuAlaLeu 384

Qy   1012 AACAACTTCGTCGATGATACATCGTACGATGTTATGAAGGAGAAAGGCGTCAACGTTATA 1071
Db    385 TyrAsnPheValSerGluArgAlaTyrAspIleLeuLysAspGlnHisPheAsnSerIle 404

Qy   1072 CCCTACCTGCGGCAATCGTGGGTTGATTTGGCGGATAAGTATATGGTAGAGGCACGGTGG 1131
Db    405 ProTyrLeuGlnArgSerTrpValSerLeuValGluGlyTyrLeuLysGluAlaTyrTrp 424

Qy   1132 TTCTACGGCGGGCACAAACCAAGTTTGGAAGAGTATTTGGAGAACTCATGGCAGTCGATA 1191
Db    425 TyrTyrAsnGlyTyrLysProSerLeuGluGluTyrLeuAsnAsnAlaLysIleSerIle 444

Qy   1192 AGTGGGCCCTGTATGTTAACGCACATATTCTTCCGAGTAACAGATTCGTTCACAAAGGAG 1251
Db    445 SerAlaProThrIleIleSerGlnLeuTyrPheThrLeuAlaAsnSerIleAspGluThr 464

Qy   1252 ACCGTCGACAGTTTGTACAAATACCACGATTTAGTTCGTTGGTCATCCTTCGTTCTGCGG 1311
Db    465 AlaIleGluSerLeuTyrGlnTyrHisAsnIleLeuTyrLeuSerGlyThrIleLeuArg 484

Qy   1312 CTTGCTGATGATTTGGGAACCTCGGTGGAAGAGGTGAGCAGAGGGGATGTGCCGAAATCA 1371
Db    485 LeuAlaAspAspLeuGlyThrSerGlnHisGluLeuGluArgGlyAspValProLysAla 504

Qy   1372 CTTCAGTGCTACATGAGTGACTACAATGCATCGGAGGCGGAGGCGCGGAAGCACGTGAAA 1431
Db    505 IleGlnCysTyrMetAsnAspThrAsnAlaSerGluArgGluAlaValGluHisValLys 524

Qy   1432 TGGCTGATAGCGGAGGTGTGGAAGAAGATGAATGCGGAGAGGGTGTCGAAGGATTCTCCA 1491
Db    525 PheLeuIleArgGluAlaTrpLysGluMetAsnThrValThrThrAlaSerAspCysPro 544

Qy   1492 TTCGGCAAAGATTTTATAGGATGTGCAGTTGATTTAGGAAGGATGGCGCAGTTGATGTAC 1551
Db    545 PheThrAspAspLeuValAlaAlaAlaAlaAsnLeuAlaArgAlaAlaGlnPheIleTyr 564

Qy   1552 CATAATGGAGATGGGCACGGCACACAACACCCTATTATACATCAACAAATGACCAGAACC 1611
Db    565 LeuAspGlyAspGlyHisGlyValGlnHisSerGluIleHisGlnGlnMetGlyGlyLeu 584

Qy   1612 TTATTCGAGCCCTTT 1626
```

```
Db        585 LeuPheGlnProTyr 589

RESULT 9
US-09-903-012-54
; Sequence 54, Application US/09903012
; Patent No. US20020094557A1
; GENERAL INFORMATION:
;  APPLICANT: Chappell, Joseph
;  APPLICANT:  No. US20020094557A11, Joseph P.
;  APPLICANT:  Starks, Courtney M.
;  APPLICANT:  Manna, Kathleen R.
;  TITLE OF INVENTION: SYNTHASES
;  FILE REFERENCE: 07678-025001
;  CURRENT APPLICATION NUMBER: US/09/903,012
;  CURRENT FILING DATE:  2001-07-11
;  PRIOR APPLICATION NUMBER: 09/398,395
;  PRIOR FILING DATE: 1999-09-17
;  PRIOR APPLICATION NUMBER: 60/100,993
;  PRIOR FILING DATE: 1998-09-18
;  PRIOR APPLICATION NUMBER: 60/130,628
;  PRIOR FILING DATE: 1999-04-22
;  PRIOR APPLICATION NUMBER: 60/150,262
;  PRIOR FILING DATE: 1999-08-23
;  NUMBER OF SEQ ID NOS: 58
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 54
;   LENGTH: 590
;   TYPE: PRT
;   ORGANISM: Salvia officinalis
US-09-903-012-54

Alignment Scores:
Pred. No.:              2.79e-136     Length:         590
Score:                  1553.50       Matches:        291
Percent Similarity:     72.48%        Conservative:   104
Best Local Similarity:  53.39%        Mismatches:     141
Query Match:            53.42%        Indels:         9
DB:                     9             Gaps:           3

US-09-938-956-6 (1-1632) x US-09-903-012-54 (1-590)

Qy          1 ATGAGACGATCCGGAAACTACAACCCTTCTCGTTGGGATGTCAACTTCATCCAATCGCTT 60
Db         51 IleArgArgSerGlyAspTyrGlnProSerLeuTrpAspPheAsnTyrIleGlnSerLeu 70

Qy         61 CTCAGTGACTATAAGGAGGACAAACACGTGATTAGGGCTTCTGAGCTGGTCACTTTGGTG 120
Db         71 AsnThrProTyrLysGluGlnArgHisPheAsnArgGlnAlaGluLeuIleMetGlnVal 90

Qy        121 AAGATGGAACTGGAGAAAGAAACGGATCAAATTCGACAACTTGAGTTGATCGATGACTTG 180
Db         91 ArgMetLeuLeuLysValLysMetGluAlaIleGlnGlnLeuGluLeuIleAspAspLeu 110

Qy        181 CAGAGGATGGGGCTGTCCGATCATTTCCAAAATGAGTTCAAAGAAATCTTGTCCTCTATA 240
Db        111 GlnTyrLeuGlyLeuSerTyrPhePheGlnAspGluIleLysGlnIleLeuSerSerIle 130

Qy        241 TATCTCGACCATCACTATTACAAGAACCCTTTTCCAAAAGAAGAAAGGGATCTCTACTCC 300
Db        131 HisAsnGluProArgTyrPheHisAsn------------------AsnAspLeuTyrPhe 144

Qy        301 ACATCTCTTGCATTTAGGCTCCTCAGAGAACATGGTTTTCAAGTCGCACAAGAGGTATTC 360
Db        145 ThrAlaLeuGlyPheArgIleLeuArgGlnHisGlyPheAsnValSerGluAspValPhe 164

Qy        361 GATAGTTTCAAGAACGAGGAG---GGTGAGTTCAAAGAAAGCCTTAGCGACGACACCAGA 417
Db        165 AspCysPheLysIleGluLysCysSerAspPheAsnAlaAsnLeuAlaGlnAspThrLys 184

Qy        418 GGATTGTTGCAACTGTATGAAGCTTCCTTTCTGTTGACGGAAGGCGAAACCACGCTCGAG 477
Db        185 GlyMetLeuGlnLeuTyrGluAlaSerPheLeuLeuArgGluGlyGluAspThrLeuGlu 204

Qy        478 TCAGCGAGGGAATTCGCCACCAAATTTTTGGAGGAAAAAGTGAACGAGGGTGGT------ 531
Db        205 LeuAlaArgArgPheSerThrArgSerLeuArgGluLysPheAspGluGlyGlyAspGlu 224

Qy        532 GTTGATGGCGACCTTTTAACAAGAATCGCATATTCTTTGGACATCCCTCTTCATTGGAGG 591
Db        225 IleAspGluAspLeuSerSerTrpIleArgHisSerLeuAspLeuProLeuHisTrpArg 244

Qy        592 ATTAAAAGGCCAAATGCACCTGTGTGGATCGAATGGTATAGGAAGAGGCCCGACATGAAT 651
Db        245 ValGlnGlyLeuGluAlaArgTrpPheLeuAspAlaTyrAlaArgArgProAspMetAsn 264

Qy        652 CCAGTAGTGTTGGAGCTTGCCATACTCGACTTAAATATTGTTCAAGCACAATTTCAAGAA 711
Db        265 ProLeuIlePheLysLeuAlaLysLeuAsnPheAsnIleValGlnAlaThrTyrGlnGlu 284

Qy        712 GAGCTCAAAGAATCCTTCAGGTGGTGGAGAAATACTGGGTTTGTTGAGAAGCTGCCCTTC 771
Db        285 GluLeuLysAspIleSerArgTrpTrpAsnSerSerCysLeuAlaGluLysLeuProPhe 304

Qy        772 GCAAGGGATAGACTGGTGGAATGCTACTTTTGGAATACTGGGATCATCGAGCCACGTCAG 831
Db        305 ValArgAspArgIleValGluCysPhePheTrpAlaIleAlaAlaPheGluProHisGln 324

Qy        832 CATGCAAGTGCAAGGATAATGATGGGCAAAGTCAACGCTCTGATTACGGTGATCGATGAT 891
Db        325 TyrSerTyrGlnArgLysMetAlaAlaValIleIleThrPheIleThrIleIleAspAsp 344

Qy        892 ATTTATGATGTCTATGGCACCTTAGAAGAACTCGAACAATTCACTGACCTCATTCGAAGA 951
Db        345 ValTyrAspValTyrGlyThrIleGluGluLeuGluLeuLeuThrAspMetIleArgArg 364

Qy        952 TGGGATATAAACTCAATCGACCAACTTCCCGATTACATGCAACTGTGCTTTCTTGCACTC 1011
Db        365 TrpAspAsnLysSerIleSerGlnLeuProTyrTyrMetGlnValCysTyrLeuAlaLeu 384

Qy       1012 AACAACTTCGTCGATGATACATCGTACGATGTTATGAAGGAGAAAGGCGTCAACGTTATA 1071
Db        385 TyrAsnPheValSerGluArgAlaTyrAspIleLeuLysAspGlnHisPheAsnSerIle 404

Qy       1072 CCCTACCTGCGGCAATCGTGGGTTGATTTGGCGGATAAGTATATGGTAGAGGCACGGTGG 1131
Db        405 ProTyrLeuGlnArgSerTrpValSerLeuValGluGlyTyrLeuLysGluAlaTyrTrp 424

Qy       1132 TTCTACGGCGGGCACAAACCAAGTTTGGAAGAGTATTTGGAGAACTCATGGCAGTCGATA 1191
Db        425 TyrTyrAsnGlyTyrLysProSerLeuGluGluTyrLeuAsnAsnAlaLysIleSerIle 444

Qy       1192 AGTGGGCCCTGTATGTTAACGCACATATTCTTCCGAGTAACAGATTCGTTCACAAAGGAG 1251
Db        445 SerAlaProThrIleIleSerGlnLeuTyrPheThrLeuAlaAsnSerIleAspGluThr 464

Qy       1252 ACCGTCGACAGTTTGTACAAATACCACGATTTAGTTCGTTGGTCATCCTTCGTTCTGCGG 1311
Db        465 AlaIleGluSerLeuTyrGlnTyrHisAsnIleLeuTyrLeuSerGlyThrIleLeuArg 484

Qy       1312 CTTGCTGATGATTTGGGAACCTCGGTGGAAGAGGTGAGCAGAGGGGATGTGCCGAAATCA 1371
Db        485 LeuAlaAspAspLeuGlyThrSerGlnHisGluLeuGluArgGlyAspValProLysAla 504

Qy       1372 CTTCAGTGCTACATGAGTGACTACAATGCATCGGAGGCGGAGGCGCGGAAGCACGTGAAA 1431
Db        505 IleGlnCysTyrMetAsnAspThrAsnAlaSerGluArgGluAlaValGluHisValLys 524

Qy       1432 TGGCTGATAGCGGAGGTGTGGAAGAAGATGAATGCGGAGAGGGTGTCGAAGGATTCTCCA 1491
Db        525 PheLeuIleArgGluAlaTrpLysGluMetAsnThrValThrThrAlaSerAspCysPro 544

Qy       1492 TTCGGCAAAGATTTTATAGGATGTGCAGTTGATTTAGGAAGGATGGCGCAGTTGATGTAC 1551
Db        545 PheThrAspAspLeuValAlaAlaAlaAlaAsnLeuAlaArgAlaAlaGlnPheIleTyr 564

Qy       1552 CATAATGGAGATGGGCACGGCACACAACACCCTATTATACATCAACAAATGACCAGAACC 1611
```

```
Db        565 LeuAspGlyAspGlyHisGlyValGlnHisSerGluIleHisGlnGlnMetGlyGlyLeu 584

Qy       1612 TTATTCGAGCCCTTT 1626

Db        585 LeuPheGlnProTyr 589

RESULT 10
US-09-900-797-54
; Sequence 54, Application US/09900797
; Publication No. US20030087406A1
; GENERAL INFORMATION:
;  APPLICANT: Chappell, Joseph
;  APPLICANT:  No. US20030087406A11, Joseph P.
;  APPLICANT:  Starks, Courtney M.
;  APPLICANT:  Manna, Kathleen R.
;  TITLE OF INVENTION: SYNTHASES
;  FILE REFERENCE: 07678-025001
;  CURRENT APPLICATION NUMBER: US/09/900,797
;  CURRENT FILING DATE:  2001-07-06
;  PRIOR APPLICATION NUMBER: US/09/398,395
;  PRIOR FILING DATE: 1999-09-17
;  PRIOR APPLICATION NUMBER: 60/130,628
;  PRIOR FILING DATE: 1999-04-22
;  PRIOR APPLICATION NUMBER: 60/150,262
;  PRIOR FILING DATE: 1999-08-23
;  NUMBER OF SEQ ID NOS: 58
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 54
;   LENGTH: 590
;   TYPE: PRT
;   ORGANISM: Salvia officinalis
US-09-900-797-54

Alignment Scores:
Pred. No.:               2.79e-136      Length:         590
Score:                   1553.50        Matches:        291
Percent Similarity:      72.48%         Conservative:   104
Best Local Similarity:   53.39%         Mismatches:     141
Query Match:             53.42%         Indels:         9
DB:                      10             Gaps:           3

US-09-938-956-6 (1-1632) x US-09-900-797-54 (1-590)

Qy          1 ATGAGACGATCCGGAAACTACAACCCTTCTCGTTGGGATGTCAACTTCATCCAATCGCTT 60

Db         51 IleArgArgSerGlyAspTyrGlnProSerLeuTrpAspPheAsnTyrIleGlnSerLeu 70

Qy         61 CTCAGTGACTATAAGGAGGACAAACACGTGATTAGGGCTTCTGAGCTGGTCACTTTGGTG 120

Db         71 AsnThrProTyrLysGluGlnArgHisPheAsnArgGlnAlaGluLeuIleMetGlnVal 90

Qy        121 AAGATGGAACTGGAGAAAGAAACGGATCAAATTCGACAACTTGAGTTGATCGATGACTTG 180

Db         91 ArgMetLeuLeuLysValLysMetGluAlaIleGlnGlnLeuGluLeuIleAspAspLeu 110

Qy        181 CAGAGGATGGGGCTGTCCGATCATTTCCAAAATGAGTTCAAAGAAATCTTGTCCTCTATA 240

Db        111 GlnTyrLeuGlyLeuSerTyrPhePheGlnAspGluIleLysGlnIleLeuSerSerIle 130

Qy        241 TATCTCGACCATCACTATTACAAGAACCCTTTTCCAAAAGAAGAAAGGGATCTCTACTCC 300

Db        131 HisAsnGluProArgTyrPheHisAsn------------------AsnAspLeuTyrPhe 144

Qy        301 ACATCTCTTGCATTTAGGCTCCTCAGAGAACATGGTTTTCAAGTCGCACAAGAGGTATTC 360

Db        145 ThrAlaLeuGlyPheArgIleLeuArgGlnHisGlyPheAsnValSerGluAspValPhe 164

Qy        361 GATAGTTTCAAGAACGAGGAG---GGTGAGTTCAAAGAAAGCCTTAGCGACGACACCAGA 417

Db        165 AspCysPheLysIleGluLysCysSerAspPheAsnAlaAsnLeuAlaGlnAspThrLys 184

Qy        418 GGATTGTTGCAACTGTATGAAGCTTCCTTTCTGTTGACGGAAGGCGAAACCACGCTCGAG 477

Db        185 GlyMetLeuGlnLeuTyrGluAlaSerPheLeuLeuArgGluGlyGluAspThrLeuGlu 204

Qy        478 TCAGCGAGGGAATTCGCCACCAAATTTTTGGAGGAAAAAGTGAACGAGGGTGGT------ 531

Db        205 LeuAlaArgArgPheSerThrArgSerLeuArgGluLysPheAspGluGlyGlyAspGlu 224

Qy        532 GTTGATGGCGACCTTTTAACAAGAATCGCATATTCTTTGGACATCCCTCTTCATTGGAGG 591

Db        225 IleAspGluAspLeuSerSerTrpIleArgHisSerLeuAspLeuProLeuHisTrpArg 244

Qy        592 ATTAAAAGGCCAAATGCACCTGTGTGGATCGAATGGTATAGGAAGAGGCCCGACATGAAT 651

Db        245 ValGlnGlyLeuGluAlaArgTrpPheLeuAspAlaTyrAlaArgArgProAspMetAsn 264

Qy        652 CCAGTAGTGTTGGAGCTTGCCATACTCGACTTAAATATTGTTCAAGCACAATTTCAAGAA 711

Db        265 ProLeuIlePheLysLeuAlaLysLeuAsnPheAsnIleValGlnAlaThrTyrGlnGlu 284

Qy        712 GAGCTCAAAGAATCCTTCAGGTGGTGGAGAAATACTGGGTTTGTTGAGAAGCTGCCCTTC 771

Db        285 GluLeuLysAspIleSerArgTrpTrpAsnSerSerCysLeuAlaGluLysLeuProPhe 304

Qy        772 GCAAGGGATAGACTGGTGGAATGCTACTTTTGGAATACTGGGATCATCGAGCCACGTCAG 831

Db        305 ValArgAspArgIleValGluCysPhePheTrpAlaIleAlaAlaPheGluProHisGln 324

Qy        832 CATGCAAGTGCAAGGATAATGATGGGCAAAGTCAACGCTCTGATTACGGTGATCGATGAT 891

Db        325 TyrSerTyrGlnArgLysMetAlaAlaValIleIleThrPheIleThrIleIleAspAsp 344

Qy        892 ATTTATGATGTCTATGGCACCTTAGAAGAACTCGAACAATTCACTGACCTCATTCGAAGA 951

Db        345 ValTyrAspValTyrGlyThrIleGluGluLeuGluLeuLeuThrAspMetIleArgArg 364

Qy        952 TGGGATATAAACTCAATCGACCAACTTCCCGATTACATGCAACTGTGCTTTCTTGCACTC 1011

Db        365 TrpAspAsnLysSerIleSerGlnLeuProTyrTyrMetGlnValCysTyrLeuAlaLeu 384

Qy       1012 AACAACTTCGTCGATGATACATCGTACGATGTTATGAAGGAGAAAGGCGTCAACGTTATA 1071

Db        385 TyrAsnPheValSerGluArgAlaTyrAspIleLeuLysAspGlnHisPheAsnSerIle 404

Qy       1072 CCCTACCTGCGGCAATCGTGGGTTGATTTGGCGGATAAGTATATGGTAGAGGCACGGTGG 1131

Db        405 ProTyrLeuGlnArgSerTrpValSerLeuValGluGlyTyrLeuLysGluAlaTyrTrp 424

Qy       1132 TTCTACGGCGGGCACAAACCAAGTTTGGAAGAGTATTTGGAGAACTCATGGCAGTCGATA 1191

Db        425 TyrTyrAsnGlyTyrLysProSerLeuGluGluTyrLeuAsnAsnAlaLysIleSerIle 444

Qy       1192 AGTGGGCCCTGTATGTTAACGCACATATTCTTCCGAGTAACAGATTCGTTCACAAAGGAG 1251

Db        445 SerAlaProThrIleIleSerGlnLeuTyrPheThrLeuAlaAsnSerIleAspGluThr 464

Qy       1252 ACCGTCGACAGTTTGTACAAATACCACGATTTAGTTCGTTGGTCATCCTTCGTTCTGCGG 1311

Db        465 AlaIleGluSerLeuTyrGlnTyrHisAsnIleLeuTyrLeuSerGlyThrIleLeuArg 484

Qy       1312 CTTGCTGATGATTTGGGAACCTCGGTGGAAGAGGTGAGCAGAGGGGATGTGCCGAAATCA 1371

Db        485 LeuAlaAspAspLeuGlyThrSerGlnHisGluLeuGluArgGlyAspValProLysAla 504

Qy       1372 CTTCAGTGCTACATGAGTGACTACAATGCATCGGAGGCGGAGGCGCGGAAGCACGTGAAA 1431

Db        505 IleGlnCysTyrMetAsnAspThrAsnAlaSerGluArgGluAlaValGluHisValLys 524

Qy       1432 TGGCTGATAGCGGAGGTGTGGAAGAAGATGAATGCGGAGAGGGTGTCGAAGGATTCTCCA 1491

Db        525 PheLeuIleArgGluAlaTrpLysGluMetAsnThrValThrThrAlaSerAspCysPro 544

Qy       1492 TTCGGCAAAGATTTTATAGGATGTGCAGTTGATTTAGGAAGGATGGCGCAGTTGATGTAC 1551
```

```
Db        545 PheThrAspAspLeuValAlaAlaAlaAlaAsnLeuAlaArgAlaAlaGlnPheIleTyr 564

Qy       1552 CATAATGGAGATGGGCACGGCACACAACACCCTATTATACATCAACAAATGACCAGAACC 1611
              :::|||||||||||||||    ||||||         ||||||||||||||
Db        565 LeuAspGlyAspGlyHisGlyValGlnHisSerGluIleHisGlnGlnMetGlyGlyLeu 584

Qy       1612 TTATTCGAGCCCTTT 1626
              ||||||:::|||:::
Db        585 LeuPheGlnProTyr 589

RESULT 11
US-09-887-586A-26
; Sequence 26, Application US/09887586A
; Patent No. US20020094556A1
; GENERAL INFORMATION:
;  APPLICANT: Chappell, Joseph
;  APPLICANT:  No. US20020094556A11, Joseph P.
;  APPLICANT:  Starks, Courtney M.
;  APPLICANT:  Manna, Kathleen R.
;  TITLE OF INVENTION: SYNTHASES
;  FILE REFERENCE: 07678-025001
;  CURRENT APPLICATION NUMBER: US/09/887,586A
;  CURRENT FILING DATE:  2001-06-22
;  PRIOR APPLICATION NUMBER: 09/398,395
;  PRIOR FILING DATE: 1999-09-17
;  PRIOR APPLICATION NUMBER: 60/130,628
;  PRIOR FILING DATE: 1999-04-22
;  PRIOR APPLICATION NUMBER: 60/150,262
;  PRIOR FILING DATE: 1999-08-23
;  NUMBER OF SEQ ID NOS: 58
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 26
;   LENGTH: 598
;   TYPE: PRT
;   ORGANISM: Salvia officinalis
US-09-887-586A-26

Alignment Scores:
Pred. No.:                2.84e-128     Length:         598
Score:                    1468.00       Matches:        281
Percent Similarity:       71.12%        Conservative:   108
Best Local Similarity:    51.37%        Mismatches:     152
Query Match:              50.48%        Indels:         6
DB:                       9             Gaps:           5

US-09-938-956-6 (1-1632) x US-09-887-586A-26 (1-598)

Qy          1 ATGAGACGATCCGGAAACTACAACCCTTCTCGTTGGGATGTCAACTTCATCCAATCGCTT 60
              :::|||||||||||||||||    |||:::   ||||||    |||:::||||||||||||
Db         54 IleArgArgSerGlyAsnTyrGlnProAlaLeuTrpAspSerAsnTyrIleGlnSerLeu 73

Qy         61 CTCAGTGACTATAAGGAGGACAAACACGTGATTAGGGCTTCTGAGCTGGTCACTTTGGTG 120
                 :::   |||   |||::::::|||:::   |||    :::||||||:::      |||
Db         74 AsnThrProTyrThrGluGluArgHisLeuAspArgLysAlaGluLeuIleValGlnVal 93

Qy        121 AAGATGGAACTGGAGAAAGAAACGGATCAAATTCGACAACTTGAGTTGATCGATGACTTG 180
              ::::::   |||:::::::::   :::   ::::::||||||||||||||     ||||||
Db         94 ArgIleLeuLeuLysGluLysMetGluProValGlnGlnLeuGluLeuIleHisAspLeu 113

Qy        181 CAGAGGATGGGGCTGTCCGATCATTTCCAAAATGAGTTCAAAGAAATCTTGTCCTCTATA 240
              :::   :::|||||||||||   |||||||:::|||   |||||||||||       |||
Db        114 LysTyrLeuGlyLeuSerAspPhePheGlnAspGluIleLysGluIleLeuGlyValIle 133

Qy        241 TATCTCGACCATCACTATTACAAGAACCCTTTTCCAAAAGAAGAAAGGGATCTCTACTCC 300
              |||   :::|||      :::   |||         |||:::   ||||||||
Db        134 TyrAsnGluHisLysCysPheHisAsnAsn---GluValGluLysMetAspLeuTyrPhe 152

Qy        301 ACATCTCTTGCATTTAGGCTCCTCAGAGAACATGGTTTTCAAGTCGCACAAGAGGTATTC 360
              |||:::|||   ||||||||||||||:::|||||||||   ::::::|||:::||||||
Db        153 ThrAlaLeuGlyPheArgLeuLeuArgGlnHisGlyPheAsnIleSerGlnAspValPhe 172

Qy        361 GATAGTTTCAAGAACGAGGAGGGT---GAGTTCAAAGAAAGCCTTAGCGACGACACCAGA 417
              :::   ||||||||||||:::|||   :::||||||   ||||||:::   ||||||:::
Db        173 AsnCysPheLysAsnGluLysGlyIleAspPheLysAlaSerLeuAlaGlnAspThrLys 192

Qy        418 GGATTGTTGCAACTGTATGAAGCTTCCTTTCTGTTGACGGAAGGCGAAACCACGCTCGAG 477
              |||:::||||||||||||||||||||||||||||||   :::||||||   |||||||||
Db        193 GlyMetLeuGlnLeuTyrGluAlaSerPheLeuLeuArgLysGlyGluAspThrLeuGlu 212

Qy        478 TCAGCGAGGGAATTCGCCACCAAATTTTTGGAGGAAAAAGTGAACGAGGGTGGT------ 531
                 ||||||||||||||||||   |||::::::|||::::::|||||||||
Db        213 LeuAlaArgGluPheAlaThrLysCysLeuGlnLysLysLeuAspGluGlyGlyAsnGlu 232

Qy        532 GTTGATGGCGACCTTTTAACAAGAATCGCATATTCTTTGGACATCCCTCTTCATTGGAGG 591
              :::|||   :::||||||   |||   :::|||||||||:::|||||||||||||||
Db        233 IleAspGluAsnLeuLeuLeuTrpIleArgHisSerLeuAspLeuProLeuHisTrpArg 252

Qy        592 ATTAAAAGGCCAAATGCACCTGTGTGGATCGAATGGTATAGGAAGAGGCCCGACATGAAT 651
              |||:::         |||      :::|||:::   |||   :::||||||||||||||
Db        253 IleGlnSerValGluAlaArgTrpPheIleAspAlaTyrAlaArgArgProAspMetAsn 272

Qy        652 CCAGTAGTGTTGGAGCTTGCCATACTCGACTTAAATATTGTTCAAGCACAATTTCAAGAA 711
              |||::::::   |||||||||   |||:::   ||||||:::||||||      |||:::
Db        273 ProLeuIlePheGluLeuAlaLysLeuAsnPheAsnIleIleGlnAlaThrHisGlnGln 292

Qy        712 GAGCTCAAAGAATCCTTCAGGTGGTGGAGAAATACTGGGTTTGTTGAGAAGCTGCCCTTC 771
              |||||||||:::      |||||||||            |||   |||||||||||||||
Db        293 GluLeuLysAspLeuSerArgTrpTrpSerArgLeuCysPheProGluLysLeuProPhe 312

Qy        772 GCAAGGGATAGACTGGTGGAATGCTACTTTTGGAATACTGGGATCATCGAGCCACGTCAG 831
                 ||||||||||||||||||   :::||||||   ||||:::   ||||||   |||
Db        313 ValArgAspArgLeuValGluSerPhePheTrpAlaValGlyMetPheGluProHisGln 332

Qy        832 CATGCAAGTGCAAGGATAATGATGGGCAAAGTCAACGCTCTGATTACGGTGATCGATGAT 891
              |||         |||   |||         :::      |||   ||||||||||||||
Db        333 HisGlyTyrGlnArgLysMetAlaAlaThrIleIleValLeuAlaThrValIleAspAsp 352

Qy        892 ATTTATGATGTCTATGGCACCTTAGAAGAACTCGAACAATTCACTGACCTCATTCGAAGA 951
              ||||||||||||||||||||||||:::||||||||||   |||||||||      :::|||
Db        353 IleTyrAspValTyrGlyThrLeuAspGluLeuGluLeuPheThrAspThrPheLysArg 372

Qy        952 TGGGATATAAACTCAATCGACCAACTTCCCGATTACATGCAACTGTGCTTTCTTGCACTC 1011
              ||||||      ||||||   :::||||||   ||||||||||||||:::      :::
Db        373 TrpAspThrGluSerIleThrArgLeuProTyrTyrMetGlnLeuCysTyrTrpGlyVal 392

Qy       1012 AACAACTTCGTCGATGATACATCGTACGATGTTATGAAGGAGAAAGGCGTCAACGTTATA 1071
              :::|||::::::   |||   :::||||||::::::||||||   |||      :::
Db        393 HisAsnTyrIleSerAspAlaAlaTyrAspIleLeuLysGluHisGlyPhePheCysLeu 412

Qy       1072 CCCTACCTGCGGCAATCGTGGGTTGATTTGGCGGATAAGTATATGGTAGAGGCACGGTGG 1131
                 |||||||||:::|||   |||||||||   :::   |||      ||||||:::|||
Db        413 GlnTyrLeuArgLysSerValValAspLeuValGluAlaTyrPheHisGluAlaLysTrp 432

Qy       1132 TTCTACGGCGGGCACAAACCAAGTTTGGAAGAGTATTTGGAGAACTCATGGCAGTCGATA 1191
              ::::::   |||:::   |||||||||:::||||||||||   :::         |||:::
Db        433 TyrHisSerGlyTyrThrProSerLeuAspGluTyrLeuAsnIleAlaLysIleSerVal 452

Qy       1192 AGTGGGCCCTGTATGTTAACGCACATATTCTTCCGAGTAACAGAT---TCGTTCACAAAG 1248
              :::   |||   :::::::::         :::|||         :::   |||
Db        453 AlaSerProAlaIleIleSerProThrTyrPheThrPheAlaAsnAlaSerHisAspThr 472

Qy       1249 GAGACCGTCGACAGTTTGTACAAATACCACGATTTAGTTCGTTGGTCATCCTTCGTTCTG 1308
                 :::|||||||||||:::|||||||||::::::         :::      :::|||
Db        473 AlaValIleAspSerLeuTyrGlnTyrHisAspIleLeuCysLeuAlaGlyIleIleLeu 492

Qy       1309 CGGCTTGCTGATGATTTGGGAACCTCGGTGGAAGAGGTGAGCAGAGGGGATGTGCCGAAA 1368
              ||||||   ||||||||||||||||||      |||::::::||||||||||||||||||
Db        493 ArgLeuProAspAspLeuGlyThrSerTyrPheGluLeuAlaArgGlyAspValProLys 512

Qy       1369 TCACTTCAGTGCTACATGAGTGACTACAATGCATCGGAGGCGGAGGCGCGGAAGCACGTG 1428
              ::::::||||||||||||   :::   ||||||||||||   ||||||   :::||||||
Db        513 ThrIleGlnCysTyrMetLysGluThrAsnAlaSerGluGluGluAlaValGluHisVal 532

Qy       1429 AAATGGCTGATAGCGGAGGTGTGGAAGAAGATGAATGCGGAGAGGGTGTCGAAGGATTCT 1488
              |||:::||||||   |||   ||||||   ||||||         ::::::
Db        533 LysPheLeuIleArgGluAlaTrpLysAspMetAsnThr---AlaIleAlaAlaGlyTyr 551
```

```
Qy       1489 CCATTCGGCAAAGATTTTATAGGATGTGCAGTTGATTTAGGAAGGATGGCGCAGTTGATG 1548
Db        552 ProPheProAspGlyMetValAlaGlyAlaAlaAsnIleGlyArgValAlaGlnPheIle 571

Qy       1549 TACCATAATGGAGATGGGCACGGCACACAACACCCTATTATACATCAACAAATGACCAGA 1608
Db        572 TyrLeuHisGlyAspGlyPheGlyValGlnHisSerLysThrTyrGluHisIleAlaGly 591

Qy       1609 ACCTTATTCGAGCCCTTTGCA 1629
Db        592 LeuLeuPheGluProTyrAla 598

RESULT 12
US-09-903-012-26
; Sequence 26, Application US/09903012
; Patent No. US20020094557A1
; GENERAL INFORMATION:
;  APPLICANT: Chappell, Joseph
;  APPLICANT:  No. US20020094557A11, Joseph P.
;  APPLICANT:  Starks, Courtney M.
;  APPLICANT:  Manna, Kathleen R.
;  TITLE OF INVENTION: SYNTHASES
;  FILE REFERENCE: 07678-025001
;  CURRENT APPLICATION NUMBER: US/09/903,012
;  CURRENT FILING DATE:  2001-07-11
;  PRIOR APPLICATION NUMBER: 09/398,395
;  PRIOR FILING DATE: 1999-09-17
;  PRIOR APPLICATION NUMBER: 60/100,993
;  PRIOR FILING DATE: 1998-09-18
;  PRIOR APPLICATION NUMBER: 60/130,628
;  PRIOR FILING DATE: 1999-04-22
;  PRIOR APPLICATION NUMBER: 60/150,262
;  PRIOR FILING DATE: 1999-08-23
;  NUMBER OF SEQ ID NOS: 58
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 26
;   LENGTH: 598
;   TYPE: PRT
;   ORGANISM: Salvia officinalis
US-09-903-012-26

Alignment Scores:
Pred. No.:              2.84e-128      Length:         598
Score:                  1468.00        Matches:        281
Percent Similarity:     71.12%         Conservative:   108
Best Local Similarity:  51.37%         Mismatches:     152
Query Match:            50.48%         Indels:         6
DB:                     9              Gaps:           5

US-09-938-956-6 (1-1632) x US-09-903-012-26 (1-598)

Qy          1 ATGAGACGATCCGGAAACTACAACCCTTCTCGTTGGATGTCAACTTCATCCAATCGCTT 60
Db         54 IleArgArgSerGlyAsnTyrGlnProAlaLeuTrpAspSerAsnTyrIleGlnSerLeu 73

Qy         61 CTCAGTGACTATAAGGAGGACAAACACGTGATTAGGGCTTCTGAGCTGGTCACTTTGGTG 120
Db         74 AsnThrProTyrThrGluGluArgHisLeuAspArgLysAlaGluLeuIleValGlnVal 93

Qy        121 AAGATGGAACTGGAGAAAGAAACGGATCAAATTCGACAACTTGAGTTGATCGATGACTTG 180
Db         94 ArgIleLeuLeuLysGluLysMetGluProValGlnGlnLeuGluLeuIleHisAspLeu 113

Qy        181 CAGAGGATGGGGCTGTCCGATCATTTCCAAAATGAGTTCAAAGAAATCTTGTCCTCTATA 240
Db        114 LysTyrLeuGlyLeuSerAspPhePheGlnAspGluIleLysGluIleLeuGlyValIle 133

Qy        241 TATCTCGACCATCACTATTACAAGAACCCTTTTCCAAAAGAAGAAAGGGATCTCTACTCC 300
Db        134 TyrAsnGluHisLysCysPheHisAsnAsn---GluValGluLysMetAspLeuTyrPhe 152

Qy        301 ACATCTCTTGCATTTAGGCTCCTCAGAGAACATGGTTTTCAAGTCGCACAAGAGGTATTC 360
Db        153 ThrAlaLeuGlyPheArgLeuLeuArgGlnHisGlyPheAsnIleSerGlnAspValPhe 172

Qy        361 GATAGTTTCAAGAACGAGGAGGGT---GAGTTCAAAGAAAGCCTTAGCGACGACACCAGA 417
Db        173 AsnCysPheLysAsnGluLysGlyIleAspPheLysAlaSerLeuAlaGlnAspThrLys 192

Qy        418 GGATTGTTGCAACTGTATGAAGCTTCCTTTCTGTTGACGGAAGGCGAAACCACGCTCGAG 477
Db        193 GlyMetLeuGlnLeuTyrGluAlaSerPheLeuLeuArgLysGlyGluAspThrLeuGlu 212

Qy        478 TCAGCGAGGGAATTCGCCACCAAATTTTTGGAGGAAAAAGTGAACGAGGGTGGT------ 531
Db        213 LeuAlaArgGluPheAlaThrLysCysLeuGlnLysLysLeuAspGluGlyGlyAsnGlu 232

Qy        532 GTTGATGGCGACCTTTTAACAAGAATCGCATATTCTTTGGACATCCCTCTTCATTGGAGG 591
Db        233 IleAspGluAsnLeuLeuLeuTrpIleArgHisSerLeuAspLeuProLeuHisTrpArg 252

Qy        592 ATTAAAAGGCCAAATGCACCTGTGTGGATCGAATGGTATAGGAAGAGGCCCGACATGAAT 651
Db        253 IleGlnSerValGluAlaArgTrpPheIleAspAlaTyrAlaArgArgProAspMetAsn 272

Qy        652 CCAGTAGTGTTGGAGCTTGCCATACTCGACTTAAATATTGTTCAAGCACAATTTCAAGAA 711
Db        273 ProLeuIlePheGluLeuAlaLysLeuAsnPheAsnIleIleGlnAlaThrHisGlnGln 292

Qy        712 GAGCTCAAAGAATCCTTCAGGTGGTGGAGAAATACTGGGTTTGTTGAGAAGCTGCCCTTC 771
Db        293 GluLeuLysAspLeuSerArgTrpTrpSerArgLeuCysPheProGluLysLeuProPhe 312

Qy        772 GCAAGGGATAGACTGGTGGAATGCTACTTTTGGAATACTGGGATCATCGAGCCACGTCAG 831
Db        313 ValArgAspArgLeuValGluSerPhePheTrpAlaValGlyMetPheGluProHisGln 332

Qy        832 CATGCAAGTGCAAGGATAATGATGGGCAAAGTCAACGCTCTGATTACGGTGATCGATGAT 891
Db        333 HisGlyTyrGlnArgLysMetAlaAlaThrIleIleValLeuAlaThrValIleAspAsp 352

Qy        892 ATTTATGATGTCTATGGCACCTTAGAAGAACTCGAACAATTCACTGACCTCATTCGAAGA 951
Db        353 IleTyrAspValTyrGlyThrLeuAspGluLeuGluLeuPheThrAspThrPheLysArg 372

Qy        952 TGGGATATAAACTCAATCGACCAACTTCCCGATTACATGCAACTGTGCTTTCTTGCACTC 1011
Db        373 TrpAspThrGluSerIleThrArgLeuProTyrTyrMetGlnLeuCysTyrTrpGlyVal 392

Qy       1012 AACAACTTCGTCGATGATACATCGTACGATGTTATGAAGGAGAAAGGCGTCAACGTTATA 1071
Db        393 HisAsnTyrIleSerAspAlaAlaTyrAspIleLeuLysGluHisGlyPhePheCysLeu 412

Qy       1072 CCCTACCTGCGGCAATCGTGGGTTGATTTGGCGGATAAGTATATGGTAGAGGCACGGTGG 1131
Db        413 GlnTyrLeuArgLysSerValValAspLeuValGluAlaTyrPheHisGluAlaLysTrp 432

Qy       1132 TTCTACGGCGGGCACAAACCAAGTTTGGAAGAGTATTTGGAGAACTCATGGCAGTCGATA 1191
Db        433 TyrHisSerGlyTyrThrProSerLeuAspGluTyrLeuAsnIleAlaLysIleSerVal 452

Qy       1192 AGTGGGCCCTGTATGTTAACGCACATATTCTTCCGAGTAACAGAT---TCGTTCACAAAG 1248
Db        453 AlaSerProAlaIleIleSerProThrTyrPheThrPheAlaAsnAlaSerHisAspThr 472

Qy       1249 GAGACCGTCGACAGTTTGTACAAATACCACGATTTAGTTCGTTGGTCATCCTTCGTTCTG 1308
Db        473 AlaValIleAspSerLeuTyrGlnTyrHisAspIleLeuCysLeuAlaGlyIleIleLeu 492

Qy       1309 CGGCTTGCTGATGATTTGGGAACCTCGGTGGAAGAGGTGAGCAGAGGGGATGTGCCGAAA 1368
Db        493 ArgLeuProAspAspLeuGlyThrSerTyrPheGluLeuAlaArgGlyAspValProLys 512

Qy       1369 TCACTTCAGTGCTACATGAGTGACTACAATGCATCGGAGGCGGAGGCGCGGAAGCACGTG 1428
```

```
Db      513 ThrIleGlnCysTyrMetLysGluThrAsnAlaSerGluGluGluAlaValGluHisVal 532

Qy     1429 AAATGGCTGATAGCGGAGGTGTGGAAGAAGATGAATGCGGAGAGGGTGTCGAAGGATTCT 1488

Db      533 LysPheLeuIleArgGluAlaTrpLysAspMetAsnThr---AlaIleAlaAlaGlyTyr 551

Qy     1489 CCATTCGGCAAAGATTTTATAGGATGTGCAGTTGATTTAGGAAGGATGGCGCAGTTGATG 1548

Db      552 ProPheProAspGlyMetValAlaGlyAlaAlaAsnIleGlyArgValAlaGlnPheIle 571

Qy     1549 TACCATAATGGAGATGGGCACGGCACACAACACCCTATTATACATCAACAAATGACCAGA 1608

Db      572 TyrLeuHisGlyAspGlyPheGlyValGlnHisSerLysThrTyrGluHisIleAlaGly 591

Qy     1609 ACCTTATTCGAGCCCTTTGCA 1629

Db      592 LeuLeuPheGluProTyrAla 598
```

RESULT 13
US-09-900-797-26
; Sequence 26, Application US/09900797
; Publication No. US20030087406A1
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. US20030087406A11, Joseph P.
; APPLICANT: Starks, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/900,797
; CURRENT FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: US/09/398,395
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 26
; LENGTH: 598
; TYPE: PRT
; ORGANISM: Salvia officinalis
US-09-900-797-26

Alignment Scores:

Pred. No.:	2.84e-128	Length:	598
Score:	1468.00	Matches:	281
Percent Similarity:	71.12%	Conservative:	108
Best Local Similarity:	51.37%	Mismatches:	152
Query Match:	50.48%	Indels:	6
DB:	10	Gaps:	5

US-09-938-956-6 (1-1632) x US-09-900-797-26 (1-598)

```
Qy        1 ATGAGACGATCCGGAAACTACAACCCTTCTCGTTGGGATGTCAACTTCATCCAATCGCTT 60

Db       54 IleArgArgSerGlyAsnTyrGlnProAlaLeuTrpAspSerAsnTyrIleGlnSerLeu 73

Qy       61 CTCAGTGACTATAAGGAGGACAAACACGTGATTAGGGCTTCTGAGCTGGTCACTTTGGTG 120

Db       74 AsnThrProTyrThrGluGluArgHisLeuAspArgLysAlaGluLeuIleValGlnVal 93

Qy      121 AAGATGGAACTGGAGAAAGAAACGGATCAAATTCGACAACTTGAGTTGATCGATGACTTG 180

Db       94 ArgIleLeuLeuLysGluLysMetGluProValGlnGlnLeuGluLeuIleHisAspLeu 113

Qy      181 CAGAGGATGGGGCTGTCCGATCATTTCCAAAATGAGTTCAAAGAAATCTTGTCCTCTATA 240

Db      114 LysTyrLeuGlyLeuSerAspPhePheGlnAspGluIleLysGluIleLeuGlyValIle 133

Qy      241 TATCTCGACCATCACTATTACAAGAACCCTTTTCCAAAAGAAGAAAGGGATCTCTACTCC 300

Db      134 TyrAsnGluHisLysCysPheHisAsnAsn---GluValGluLysMetAspLeuTyrPhe 152

Qy      301 ACATCTCTTGCATTTAGGCTCCTCAGAGAACATGGTTTTCAAGTCGCACAAGAGGTATTC 360

Db      153 ThrAlaLeuGlyPheArgLeuLeuArgGlnHisGlyPheAsnIleSerGlnAspValPhe 172

Qy      361 GATAGTTTCAAGAACGAGGAGGGT---GAGTTCAAAGAAAGCCTTAGCGACGACACCAGA 417

Db      173 AsnCysPheLysAsnGluLysGlyIleAspPheLysAlaSerLeuAlaGlnAspThrLys 192

Qy      418 GGATTGTTGCAACTGTATGAAGCTTCCTTTCTGTTGACGGAAGGCGAAACCACGCTCGAG 477

Db      193 GlyMetLeuGlnLeuTyrGluAlaSerPheLeuLeuArgLysGlyGluAspThrLeuGlu 212

Qy      478 TCAGCGAGGGAATTCGCCACCAAATTTTTGGAGGAAAAAGTGAACGAGGGTGGT------ 531

Db      213 LeuAlaArgGluPheAlaThrLysCysLeuGlnLysLysLeuAspGluGlyGlyAsnGlu 232

Qy      532 GTTGATGGCGACCTTTTAACAAGAATCGCATATTCTTTGGACATCCCTCTTCATTGGAGG 591

Db      233 IleAspGluAsnLeuLeuLeuTrpIleArgHisSerLeuAspLeuProLeuHisTrpArg 252

Qy      592 ATTAAAAGGCCAAATGCACCTGTGTGGATCGAATGGTATAGGAAGAGGCCCGACATGAAT 651

Db      253 IleGlnSerValGluAlaArgTrpPheIleAspAlaTyrAlaArgArgProAspMetAsn 272

Qy      652 CCAGTAGTGTTGGAGCTTGCCATACTCGACTTAAATATTGTTCAAGCACAATTTCAAGAA 711

Db      273 ProLeuIlePheGluLeuAlaLysLeuAsnPheAsnIleIleGlnAlaThrHisGlnGln 292

Qy      712 GAGCTCAAAGAATCCTTCAGGTGGTGGAGAAATACTGGGTTTGTTGAGAAGCTGCCCTTC 771

Db      293 GluLeuLysAspLeuSerArgTrpTrpSerArgLeuCysPheProGluLysLeuProPhe 312

Qy      772 GCAAGGGATAGACTGGTGGAATGCTACTTTTGGAATACTGGGATCATCGAGCCACGTCAG 831

Db      313 ValArgAspArgLeuValGluSerPhePheTrpAlaValGlyMetPheGluProHisGln 332

Qy      832 CATGCAAGTGCAAGGATAATGATGGGCAAAGTCAACGCTCTGATTACGGTGATCGATGAT 891

Db      333 HisGlyTyrGlnArgLysMetAlaAlaThrIleIleValLeuAlaThrValIleAspAsp 352

Qy      892 ATTTATGATGTCTATGGCACCTTAGAAGAACTCGAACAATTCACTGACCTCATTCGAAGA 951

Db      353 IleTyrAspValTyrGlyThrLeuAspGluLeuGluLeuPheThrAspThrPheLysArg 372

Qy      952 TGGGATATAAACTCAATCGACCAACTTCCCGATTACATGCAACTGTGCTTTCTTGCACTC 1011

Db      373 TrpAspThrGluSerIleThrArgLeuProTyrTyrMetGlnLeuCysTyrTrpGlyVal 392

Qy     1012 AACAACTTCGTCGATGATACATCGTACGATGTTATGAAGGAGAAAGGCGTCAACGTTATA 1071

Db      393 HisAsnTyrIleSerAspAlaAlaTyrAspIleLeuLysGluHisGlyPhePheCysLeu 412

Qy     1072 CCCTACCTGCGGCAATCGTGGGTTGATTTGGCGGATAAGTATATGGTAGAGGCACGGTGG 1131

Db      413 GlnTyrLeuArgLysSerValValAspLeuValGluAlaTyrPheHisGluAlaLysTrp 432

Qy     1132 TTCTACGGCGGGCACAAACCAAGTTTGGAAGAGTATTTGGAGAACTCATGGCAGTCGATA 1191

Db      433 TyrHisSerGlyTyrThrProSerLeuAspGluTyrLeuAsnIleAlaLysIleSerVal 452

Qy     1192 AGTGGGCCCTGTATGTTAACGCACATATTCTTCCGAGTAACAGAT---TCGTTCACAAAG 1248

Db      453 AlaSerProAlaIleIleSerProThrTyrPheThrPheAlaAsnAlaSerHisAspThr 472

Qy     1249 GAGACCGTCGACAGTTTGTACAAATACCACGATTTAGTTCGTTGGTCATCCTTCGTTCTG 1308

Db      473 AlaValIleAspSerLeuTyrGlnTyrHisAspIleLeuCysLeuAlaGlyIleIleLeu 492

Qy     1309 CGGCTTGCTGATGATTTGGGAACCTCGGTGGAAGAGGTGAGCAGAGGGGATGTGCCGAAA 1368

Db      493 ArgLeuProAspAspLeuGlyThrSerTyrPheGluLeuAlaArgGlyAspValProLys 512
```

```
Qy       1369 TCACTTCAGTGCTACATGAGTGACTACAATGCATCGGAGGCGGAGGCGCGGAAGCACGTG 1428
              :::::|||||||||||   :::   ||||||||||   |||||    :::||||||
Db        513 ThrIleGlnCysTyrMetLysGluThrAsnAlaSerGluGluGluAlaValGluHisVal 532

Qy       1429 AAATGGCTGATAGCGGAGGTGTGGAAGAAGATGAATGCGGAGAGGGTGTCGAAGGATTCT 1488
              |||:::|||||   |||   ||||||   ||||||        ::::::
Db        533 LysPheLeuIleArgGluAlaTrpLysAspMetAsnThr---AlaIleAlaAlaGlyTyr 551

Qy       1489 CCATTCGGCAAAGATTTTATAGGATGTGCAGTTGATTTAGGAAGGATGGCGCAGTTGATG 1548
              ||||||         :::        |||   ::::::||||||:::||||||   :::
Db        552 ProPheProAspGlyMetValAlaGlyAlaAlaAsnIleGlyArgValAlaGlnPheIle 571

Qy       1549 TACCATAATGGAGATGGGCACGGCACACAACACCCTATTATACATCAACAAATGACCAGA 1608
              |||   :::|||||||||   |||   ||||||        ::::::   :::
Db        572 TyrLeuHisGlyAspGlyPheGlyValGlnHisSerLysThrTyrGluHisIleAlaGly 591

Qy       1609 ACCTTATTCGAGCCCTTTGCA 1629
                 |||||||||||:::|||
Db        592 LeuLeuPheGluProTyrAla 598
```

RESULT 14
US-09-887-586A-32
; Sequence 32, Application US/09887586A
; Patent No. US20020094556A1
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. US20020094556A11, Joseph P.
; APPLICANT: Starks, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/887,586A
; CURRENT FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: 09/398,395
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 32
; LENGTH: 556
; TYPE: PRT
; ORGANISM: Solanum tuberosum
US-09-887-586A-32

```
Alignment Scores:
Pred. No.:               4.13e-69      Length:         556
Score:                   836.00        Matches:        182
Percent Similarity:      54.68%        Conservative:   116
Best Local Similarity:   33.39%        Mismatches:     207
Query Match:             28.75%        Indels:         40
DB:                      9             Gaps:           10
```

US-09-938-956-6 (1-1632) x US-09-887-586A-32 (1-556)

```
Qy          1 ATGAGACGATCCGGAAACTACAACCCTTCTCGTTGGGATGTCAACTTCATCCAATCGCTT 60
              :::|||         :::::::::||||||   |||            |||
Db         17 ValArgProIleAlaAspPheSerProSerLeuTrpGlyAspArgPheHisSerPheSer 36

Qy         61 CTCAGTGACTATAAGGAGGACAAACACGTGATTAGGGCTTCTGAGCTGGTCACTTTGGTG 120
              |||   :::            |||:::            |||   |||:::   ||||||
Db         37 LeuAspAsnGlnIleAlaGlyLysTyr---------AlaGlnGluIleGluThrLeu--- 52

Qy        121 AAGATGGAACTGGAGAAAGAAACGGATCAAATT--------------------------- 153
                                   ||||||   :::|||
Db         53 ---------------LysGluGlnSerArgIleIleLeuSerAlaSerSerArgArgThr 67

Qy        154 ------CGACAACTTGAGTTGATCGATGACTTGCAGAGGATGGGGCTGTCCGATCATTTC 207
                    :::|||:::|||||||||   ::::::|||:::|||::::::   ||||||
Db         68 LeuAlaGluLysLeuAspLeuIleAspIleValGluArgLeuGlyIleAlaTyrHisPhe 87

Qy        208 CAAAATGAGTTCAAAGAAATCTTGTCCTCTATATAT---------------CTCGACCAT 252
              :::   :::      ::::::|||            |||                :::
Db         88 GluLysGlnIleAspAspMetLeuAspGlnPheTyrLysAlaAspProAsnPheGluAla 107

Qy        253 CACTATTACAAGAACCCTTTTCCAAAAGAAGAAAGGGATCTCTACTCCACATCTCTTGCA 312
              |||   |||                              ||||||   :::   |||:::
Db        108 HisGluTyr------------------------AsnAspLeuGlnThrLeuSerValGln 119

Qy        313 TTTAGGCTCCTCAGAGAACATGGTTTTCAAGTCGCACAAGAGGTATTCGATAGTTTCAAG 372
              |||||||||||||||:::||||||:::   ::::::   ::::::|||      |||:::
Db        120 PheArgLeuLeuArgGlnHisGlyTyrAsnIleSerProLysLeuPheIleArgPheGln 139

Qy        373 AACGAGGAGGGTGAGTTCAAAGAAAGCCTTAGCGACGACACCAGAGGATTGTTGCAACTG 432
              :::   :::|||:::||||||||||||||||||   :::|||   :::|||||||||   |||
Db        140 AspAlaLysGlyLysPheLysGluSerLeuCysAsnAspIleLysGlyLeuLeuAsnLeu 159

Qy        433 TATGAAGCTTCCTTTCTGTTGACGGAAGGCGAAACCACGCTCGAGTCAGCGAGGGAATTC 492
              ||||||||||||   :::   |||   ||||||      ||||||   |||      |||
Db        160 TyrGluAlaSerHisValArgThrHisGlyGluAspIleLeuGluGluAlaLeuAlaPhe 179

Qy        493 GCCACCAAATTTTTGGAGGAAAAAGTGAACGAGGGTGGTGTTGATGGCGACCTTTTAACA 552
              :::|||      ||||||                    :::         |||
Db        180 SerThrAlaHisLeuGluSerAlaAlaProHis------LeuLysSerProLeuSerLys 197

Qy        553 AGAATCGCATATTCTTTGGACATCCCTCTTCATTGGAGGATTAAAAGGCCAAATGCACCT 612
              ::::::   ::::::|||:::            ||||||      |||   |||
Db        198 GlnValThrHisAlaLeuGluGlnSerLeuHisLysSerIleProArgValGluThrArg 217

Qy        613 GTGTGGATCGAATGGTATAGGAAGAGGCCCGACATGAATCCAGTAGTGTTGGAGCTTGCC 672
                 :::|||      |||   :::         |||   |||:::|||:::      |||
Db        218 TyrPheIleSerIleTyrGluGluGluGluGlnLysAsnAspValLeuLeuGlnPheAla 237

Qy        673 ATACTCGACTTAAATATTGTTCAAGCACAATTTCAAGAAGAGCTCAAAGAATCCTTCAGG 732
                 ||||||   |||::::::|||         ::::::||||||   |||      |||
Db        238 LysLeuAspPheAsnLeuLeuGlnMetLeuHisLysGlnGluLeuSerGluValSerArg 257

Qy        733 TGGTGGAGAAATACTGGGTTTGTTGAGAAGCTGCCCTTCGCAAGGGATAGACTGGTGGAA 792
              ||||||::::::         ||||||      ||||||:::||||||||||||   ||||||
Db        258 TrpTrpLysAspLeuAspPheValThrThrLeuProTyrAlaArgAspArgAlaValGlu 277

Qy        793 TGCTACTTTTGGAATACTGGGATCATCGAGCCACGTCAGCATGCAAGTGCAAGGATAATG 852
              ||||||||||||         |||:::         |||::::::   ||||||:::|||
Db        278 CysTyrPheTrpThrMetGlyValTyrAlaGluProGlnTyrSerGlnAlaArgValMet 297

Qy        853 ATGGGCAAAGTCAACGCTCTGATTACGGTGATCGATGATATTTATGATGTCTATGGCACC 912
              :::   |||      |||:::|||:::::::::||||||   :::|||   ||||||
Db        298 LeuAlaLysThrIleAlaMetIleSerIleValAspAspThrPheAspAlaTyrGlyIle 317

Qy        913 TTAGAAGAACTCGAACAATTCACTGACCTCATTCGAAGATGGGATATAAACTCAATCGAC 972
              ::::::|||||||||   :::||||||   |||:::||||||||||||:::   ||||||
Db        318 ValLysGluLeuGluIleTyrThrAspAlaIleGlnArgTrpAspIleSerGlnIleAsp 337

Qy        973 CAACTTCCCGATTACATGCAACTGTGCTTTCTTGCACTCAACAACTTCGTCGATGATACA 1032
              :::||||||||||||||||:::::: :::      ||||||   :::      :::|||
Db        338 ArgLeuProAspTyrMetLysIleSerTyrLysAlaLeuLeuAspLeuTyrAsnAspTyr 357

Qy       1033 TCGTACGATGTTATGAAGGAGAAAGGCGTCAACGTTATACCCTACCTGCGGCAATCGTGG 1092
                 ::::::   |||:::            :::|||:::   |||   ::::::
Db        358 GluMetGluLeuSerLysAspGlyArgSerAspValValHisTyrAlaLysGluArgMet 377

Qy       1093 GTTGATTTGGCGGATAAGTATATGGTAGAGGCACGGTGGTTCTACGGCGGGCACAAACCA 1152
                 ::::::            |||   ||||||||:::||||||      |||:::   |||
Db        378 LysGluIleValArgAsnTyrPheValGluAlaLysTrpPheIleGluGlyTyrMetPro 397

Qy       1153 AGTTTGGAAGAGTATTTGGAGAACTCATGGCAGTCGATAAGTGGGCCCTGTATGTTAACG 1212
              :::   |||||||||   |||:::      :::   |||            :::||||||
Db        398 ProValSerGluTyrLeuSerAsnAla---LeuAlaThrSerThrTyrTyrLeuLeuThr 416

Qy       1213 CACATATTCTTCCGAGTAACAGATTCGTTCACAAAGGAGACCGTCGACAGTTTGTACAAA 1272
                 :::            |||      |||:::      :::   |||      |||
Db        417 ThrThrSerTyrLeuGlyMetLysSerAlaAsnLysGlnAspPheGluTrpLeuAlaLys 436
```

```
Qy       1273 TACCACGATTTAGTTCGTTGGTCATCCTTCGTTCTGCGGCTTGCTGATGATTTGGGAACC 1332
Db        437 AsnProLysIleLeuGluAlaAsnValThrLeuCysArgValIleAspAspIleAlaThr 456
Qy       1333 TCGGTGGAAGAGGTGAGCAGAGGGGATGTGCCGAAATCACTTCAGTGCTACATGAGTGAC 1392
Db        457 TyrGluValGluLysGlyArgGlyGlnIleAlaThrGlyIleGluCysTyrMetArgAsp 476
Qy       1393 TACAATGCATCGGAGGCGGAGGCGCGGAAGCACGTGAAATGGCTGATAGCGGAGGTGTGG 1452
Db        477 TyrGlyValSerThrGluLysAlaMetGluLysPheGlnGluMetAlaGluThrAlaTrp 496
Qy       1453 AAGAAGATGAATGCGGAGAGGGTGTCGAAGGATTCTCCATTCGGCAAAGATTTTATAGGA 1512
Db        497 LysAspValAsn---GluGlyIleLeuArgProThrProValSerThrGluIleLeuThr 515
Qy       1513 TGTGCAGTTGATTTAGGAAGGATGGCGCAGTTGATGTAC---CATAATGGAGATGGGCAC 1569
Db        516 ArgIleLeuAsnLeuAlaArgIleIleAspValThrTyrLysHisAsnGlnAspGlyTyr 535
Qy       1570 GGCACACAACACCCT 1584
Db        536 ------ThrHisPro 538
```

RESULT 15
US-09-903-012-32
; Sequence 32, Application US/09903012
; Patent No. US20020094557A1
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. US20020094557A11, Joseph P.
; APPLICANT: Starks, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/903,012
; CURRENT FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: 09/398,395
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: 60/100,993
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 32
; LENGTH: 556
; TYPE: PRT
; ORGANISM: Solanum tuberosum
US-09-903-012-32

Alignment Scores:

Pred. No.:	4.13e-69	Length:	556
Score:	836.00	Matches:	182
Percent Similarity:	54.68%	Conservative:	116
Best Local Similarity:	33.39%	Mismatches:	207
Query Match:	28.75%	Indels:	40
DB:	9	Gaps:	10

US-09-938-956-6 (1-1632) x US-09-903-012-32 (1-556)

```
Qy          1 ATGAGACGATCCGGAAACTACAACCCTTCTCGTTGGGATGTCAACTTCATCCAATCGCTT 60
Db         17 ValArgProIleAlaAspPheSerProSerLeuTrpGlyAspArgPheHisSerPheSer 36
Qy         61 CTCAGTGACTATAAGGAGGACAAACACGTGATTAGGGCTTCTGAGCTGGTCACTTTGGTG 120
Db         37 LeuAspAsnGlnIleAlaGlyLysTyr---------AlaGlnGluIleGluThrLeu--- 52
Qy        121 AAGATGGAACTGGAGAAAGAAACGGATCAAATT--------------------------- 153
Db         53 ---------------LysGluGlnSerArgIleIleLeuSerAlaSerSerArgArgThr 67
Qy        154 ------CGACAACTTGAGTTGATCGATGACTTGCAGAGGATGGGGCTGTCCGATCATTTC 207
Db         68 LeuAlaGluLysLeuAspLeuIleAspIleValGluArgLeuGlyIleAlaTyrHisPhe 87
Qy        208 CAAAATGAGTTCAAAGAAATCTTGTCCTCTATATAT---------------CTCGACCAT 252
Db         88 GluLysGlnIleAspAspMetLeuAspGlnPheTyrLysAlaAspProAsnPheGluAla 107
Qy        253 CACTATTACAAGAACCCTTTTCCAAAAGAAGAAAGGGATCTCTACTCCACATCTCTTGCA 312
Db        108 HisGluTyr-----------------------AsnAspLeuGlnThrLeuSerValGln 119
Qy        313 TTTAGGCTCCTCAGAGAACATGGTTTTCAAGTCGCACAAGAGGTATTCGATAGTTTCAAG 372
Db        120 PheArgLeuLeuArgGlnHisGlyTyrAsnIleSerProLysLeuPheIleArgPheGln 139
Qy        373 AACGAGGAGGGTGAGTTCAAAGAAAGCCTTAGCGACGACACCAGAGGATTGTTGCAACTG 432
Db        140 AspAlaLysGlyLysPheLysGluSerLeuCysAsnAspIleLysGlyLeuLeuAsnLeu 159
Qy        433 TATGAAGCTTCCTTTCTGTTGACGGAAGGCGAAACCACGCTCGAGTCAGCGAGGGAATTC 492
Db        160 TyrGluAlaSerHisValArgThrHisGlyGluAspIleLeuGluGluAlaLeuAlaPhe 179
Qy        493 GCCACCAAATTTTTGGAGGAAAAAGTGAACGAGGGTGGTGTTGATGGCGACCTTTTAACA 552
Db        180 SerThrAlaHisLeuGluSerAlaAlaProHis------LeuLysSerProLeuSerLys 197
Qy        553 AGAATCGCATATTCTTTGGACATCCCTCTTCATTGGAGGATTAAAAGGCCAAATGCACCT 612
Db        198 GlnValThrHisAlaLeuGluGlnSerLeuHisLysSerIleProArgValGluThrArg 217
Qy        613 GTGTGGATCGAATGGTATAGGAAGAGGCCCGACATGAATCCAGTAGTGTTGGAGCTTGCC 672
Db        218 TyrPheIleSerIleTyrGluGluGluGluGlnLysAsnAspValLeuLeuGlnPheAla 237
Qy        673 ATACTCGACTTAAATATTGTTCAAGCACAATTTCAAGAAGAGCTCAAAGAATCCTTCAGG 732
Db        238 LysLeuAspPheAsnLeuLeuGlnMetLeuHisLysGlnGluLeuSerGluValSerArg 257
Qy        733 TGGTGGAGAAATACTGGGTTTGTTGAGAAGCTGCCCTTCGCAAGGGATAGACTGGTGGAA 792
Db        258 TrpTrpLysAspLeuAspPheValThrThrLeuProTyrAlaArgAspArgAlaValGlu 277
Qy        793 TGCTACTTTTGGAATACTGGGATCATCGAGCCACGTCAGCATGCAAGTGCAAGGATAATG 852
Db        278 CysTyrPheTrpThrMetGlyValTyrAlaGluProGlnTyrSerGlnAlaArgValMet 297
Qy        853 ATGGGCAAAGTCAACGCTCTGATTACGGTGATCGATGATATTTATGATGTCTATGGCACC 912
Db        298 LeuAlaLysThrIleAlaMetIleSerIleValAspAspThrPheAspAlaTyrGlyIle 317
Qy        913 TTAGAAGAACTCGAACAATTCACTGACCTCATTCGAAGATGGGATATAAACTCAATCGAC 972
Db        318 ValLysGluLeuGluIleTyrThrAspAlaIleGlnArgTrpAspIleSerGlnIleAsp 337
Qy        973 CAACTTCCCGATTACATGCAACTGTGCTTTCTTGCACTCAACAACTTCGTCGATGATACA 1032
Db        338 ArgLeuProAspTyrMetLysIleSerTyrLysAlaLeuLeuAspLeuTyrAsnAspTyr 357
Qy       1033 TCGTACGATGTTATGAAGGAGAAAGGCGTCAACGTTATACCCTACCTGCGGCAATCGTGG 1092
Db        358 GluMetGluLeuSerLysAspGlyArgSerAspValValHisTyrAlaLysGluArgMet 377
Qy       1093 GTTGATTTGGCGGATAAGTATATGGTAGAGGCACGGTGGTTCTACGGCGGGCACAAACCA 1152
Db        378 LysGluIleValArgAsnTyrPheValGluAlaLysTrpPheIleGluGlyTyrMetPro 397
Qy       1153 AGTTTGGAAGAGTATTTGGAGAACTCATGGCAGTCGATAAGTGGGCCCTGTATGTTAACG 1212
Db        398 ProValSerGluTyrLeuSerAsnAla---LeuAlaThrSerThrTyrTyrLeuLeuThr 416
```

```
Qy      1213 CACATATTCTTCCGAGTAACAGATTCGTTCACAAAGGAGACCGTCGACAGTTTGTACAAA 1272
                      :::            |||      |||:::      :::  |||   |||
Db       417 ThrThrSerTyrLeuGlyMetLysSerAlaAsnLysGlnAspPheGluTrpLeuAlaLys 436

Qy      1273 TACCACGATTTAGTTCGTTGGTCATCCTTCGTTCTGCGGCTTGCTGATGATTTGGGAACC 1332
                      ::::::      :::      :::   |||:::   ||||||:::  |||
Db       437 AsnProLysIleLeuGluAlaAsnValThrLeuCysArgValIleAspAspIleAlaThr 456

Qy      1333 TCGGTGGAAGAGGTGAGCAGAGGGGATGTGCCGAAATCACTTCAGTGCTACATGAGTGAC 1392
                      |||      ||||||   :::         ::::::|||||||||  |||
Db       457 TyrGluValGluLysGlyArgGlyGlnIleAlaThrGlyIleGluCysTyrMetArgAsp 476

Qy      1393 TACAATGCATCGGAGGCGGAGGCGCGGAAGCACGTGAAATGGCTGATAGCGGAGGTGTGG 1452
             |||      |||      :::|||   :::      :::   :::           |||
Db       477 TyrGlyValSerThrGluLysAlaMetGluLysPheGlnGluMetAlaGluThrAlaTrp 496

Qy      1453 AAGAAGATGAATGCGGAGAGGGTGTCGAAGGATTCTCCATTCGGCAAAGATTTTATAGGA 1512
             |||   :::|||   |||   :::   :::   :::|||         :::  :::
Db       497 LysAspValAsn---GluGlyIleLeuArgProThrProValSerThrGluIleLeuThr 515

Qy      1513 TGTGCAGTTGATTTAGGAAGGATGGCGCAGTTGATGTAC---CATAATGGAGATGGGCAC 1569
                   ::::::|||   |||:::      :::   |||   ||||||  |||||||:::
Db       516 ArgIleLeuAsnLeuAlaArgIleIleAspValThrTyrLysHisAsnGlnAspGlyTyr 535

Qy      1570 GGCACACAACACCCT 1584
                      ||||||
Db       536 ------ThrHisPro 538
```

Search completed: February 25, 2004, 06:03:40
Job time : 113 secs

GenCore version 5.1.6

OM nucleic - protein search, using frame_plus_n2p model

```
Run on:         February 25, 2004, 05:40:59 ; Search time 32 Seconds
                                  (without alignments)
                                  9811.536 Million cell updates/sec

Title:          US-09-938-956-6
Perfect score:  2908
Sequence:       1 atgagacgatccggaaacta..........tattcgagccctttgcatga 1632

Scoring table:  BLOSUM62
                Xgapop 10.0 , Xgapext  0.5
                Ygapop 10.0 , Ygapext  0.5
                Fgapop  6.0 , Fgapext  7.0
                Delop   6.0 , Delext   7.0

Searched:       283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters:      566732

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US09938956/runat_24022004_144701_3078/app_query.fasta_1.1799
-DB=PIR_78 -QFMT=fastan -SUFFIX=rpr -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09938956_@CGN_1_1_43_@runat_24022004_144701_3078 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :      PIR_78:*
                1:  pir1:*
                2:  pir2:*
                3:  pir3:*
                4:  pir4:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	2876	98.9	599	2	A48863	limonene cyclase -
	2	1147	39.4	591	2	H84633	probable limonene
	3	1084.5	37.3	600	2	F71434	probable limonene
	4	1039	35.7	1024	2	G71434	probable limonene
	5	822.5	28.3	559	2	T08174	sesquiterpene cycl
	6	812	27.9	520	2	A56118	vetispiradiene syn
	7	809.5	27.8	550	2	T03714	5-epi-aristolochen
	8	796	27.4	554	2	S68366	(+)-delta-cadinene
	9	776	26.7	554	2	S68365	(+)-delta-cadinene
	10	734	25.2	548	2	T06266	germacrene C synth
	11	724	24.9	548	2	T06265	germacrene C synth
	12	623	21.4	632	2	E96723	hypothetical prote
	13	600	20.6	608	2	G86443	probable terpene s
	14	565	19.4	421	2	C96642	hypothetical prote
	15	562.5	19.3	350	2	C56118	vetispiradiene syn
	16	552.5	19.0	530	2	G96588	hypothetical prote
	17	545.5	18.8	598	2	H86460	hypothetical prote
	18	530.5	18.2	573	2	T05328	hypothetical prote
	19	515.5	17.7	300	2	B56118	vetispiradiene syn
	20	510.5	17.6	598	2	F96684	probable terpene s
	21	508	17.5	535	2	T06285	hypothetical prote
	22	498.5	17.1	582	2	C71424	hypothetical prote
	23	484	16.6	471	2	T06287	hypothetical prote
	24	479	16.5	598	2	T00509	probable vetispira
	25	457	15.7	612	2	T05331	hypothetical prote
	26	429.5	14.8	383	2	D71424	hypothetical prote
	27	414	14.2	785	2	G96825	hypothetical prote
	28	414	14.2	785	2	T52059	ent-kaurene syntha
	29	405	13.9	598	2	T05329	hypothetical prote
	30	359	12.3	789	2	T09672	ent-kaurene syntha
	31	353.5	12.2	870	2	A96637	hypothetical prote
	32	309.5	10.6	802	2	D85035	ent-kaurene synthe
	33	275	9.5	823	2	T02959	kaurene synthase A
	34	272.5	9.4	801	2	T06783	ent-kaurene syntha
c	35	138.5	4.8	1169	2	S38181	flocculation prote
	36	138	4.7	203	2	H96525	probable terpene c
	37	137.5	4.7	148	2	T03982	5-epi-aristolochen
	38	130.5	4.5	2829	2	A42771	reticulocyte-bindi
	39	130	4.5	850	2	G70332	conserved hypothet
c	40	129.5	4.5	534	2	T39903	serine-rich protei
	41	127	4.4	686	2	T10684	hypothetical prote
	42	125.5	4.3	1030	2	S73944	hypothetical prote
	43	124	4.3	1939	2	I48175	myosin heavy chain
	44	123	4.2	1713	2	A55347	adhesive ligand ep
	45	123	4.2	2672	2	A48126	translation activa

ALIGNMENTS

```
RESULT 1
A48863
limonene cyclase - spearmint
C;Species: Mentha spicata (spearmint)
C;Date: 12-May-1995 #sequence_revision 19-May-1995 #text_change 05-Nov-1999
C;Accession: A48863
R;Colby, S.M.; Alonso, W.R.; Katahira, E.J.; McGarvey, D.J.; Croteau, R.
J. Biol. Chem. 268, 23016-23024, 1993
A;Title: 4S-limonene synthase from the oil glands of spearmint (Mentha spicata). cDNA i
A;Reference number: A48863; MUID:94043077; PMID:8226816
A;Accession: A48863
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-599 <COL>
A;Cross-references: GB:L13459; NID:g410229; PIDN:AAC37366.1; PID:g410230
C;Superfamily: vetispiradiene synthase 1

Alignment Scores:
Pred. No.:              4.87e-204    Length:        599
Score:                  2876.00      Matches:       542
Percent Similarity:     100.00%      Conservative:  0
Best Local Similarity:  100.00%      Mismatches:    0
Query Match:            98.90%       Indels:        0
DB:                     2            Gaps:          0

US-09-938-956-6 (1-1632) x A48863 (1-599)

Qy          4 AGACGATCCGGAAACTACAACCCTTCTCGTTGGGATGTCAACTTCATCCAATCGCTTCTC 63
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         58 ArgArgSerGlyAsnTyrAsnProSerArgTrpAspValAsnPheIleGlnSerLeuLeu 77

Qy         64 AGTGACTATAAGGAGGACAAACACGTGATTAGGGCTTCTGAGCTGGTCACTTTGGTGAAG 123
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         78 SerAspTyrLysGluAspLysHisValIleArgAlaSerGluLeuValThrLeuValLys 97

Qy        124 ATGGAACTGGAGAAAGAAACGGATCAAATTCGACAACTTGAGTTGATCGATGACTTGCAG 183
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         98 MetGluLeuGluLysGluThrAspGlnIleArgGlnLeuGluLeuIleAspAspLeuGln 117
```

```
Qy        184 AGGATGGGGCTGTCCGATCATTTCCAAAATGAGTTCAAAGAAATCTTGTCCTCTATATAT 243
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        118 ArgMetGlyLeuSerAspHisPheGlnAsnGluPheLysGluIleLeuSerSerIleTyr 137

Qy        244 CTCGACCATCACTATTACAAGAACCCTTTTCCAAAAGAAGAAAGGGATCTCTACTCCACA 303
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        138 LeuAspHisHisTyrTyrLysAsnProPheProLysGluGluArgAspLeuTyrSerThr 157

Qy        304 TCTCTTGCATTTAGGCTCCTCAGAGAACATGGTTTTCAAGTCGCACAAGAGGTATTCGAT 363
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        158 SerLeuAlaPheArgLeuLeuArgGluHisGlyPheGlnValAlaGlnGluValPheAsp 177

Qy        364 AGTTTCAAGAACGAGGAGGGTGAGTTCAAAGAAAGCCTTAGCGACGACACCAGAGGATTG 423
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        178 SerPheLysAsnGluGluGlyGluPheLysGluSerLeuSerAspAspThrArgGlyLeu 197

Qy        424 TTGCAACTGTATGAAGCTTCCTTTCTGTTGACGGAAGGCGAAACCACGCTCGAGTCAGCG 483
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        198 LeuGlnLeuTyrGluAlaSerPheLeuLeuThrGluGlyGluThrThrLeuGluSerAla 217

Qy        484 AGGGAATTCGCCACCAAATTTTTGGAGGAAAAAGTGAACGAGGGTGGTGTTGATGGCGAC 543
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        218 ArgGluPheAlaThrLysPheLeuGluGluLysValAsnGluGlyGlyValAspGlyAsp 237

Qy        544 CTTTTAACAAGAATCGCATATTCTTTGGACATCCCTCTTCATTGGAGGATTAAAAGGCCA 603
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        238 LeuLeuThrArgIleAlaTyrSerLeuAspIleProLeuHisTrpArgIleLysArgPro 257

Qy        604 AATGCACCTGTGTGGATCGAATGGTATAGGAAGAGGCCCGACATGAATCCAGTAGTGTTG 663
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        258 AsnAlaProValTrpIleGluTrpTyrArgLysArgProAspMetAsnProValValLeu 277

Qy        664 GAGCTTGCCATACTCGACTTAAATATTGTTCAAGCACAATTTCAAGAAGAGCTCAAAGAA 723
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        278 GluLeuAlaIleLeuAspLeuAsnIleValGlnAlaGlnPheGlnGluGluLeuLysGlu 297

Qy        724 TCCTTCAGGTGGTGGAGAAATACTGGGTTTGTTGAGAAGCTGCCCTTCGCAAGGGATAGA 783
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        298 SerPheArgTrpTrpArgAsnThrGlyPheValGluLysLeuProPheAlaArgAspArg 317

Qy        784 CTGGTGGAATGCTACTTTTGGAATACTGGGATCATCGAGCCACGTCAGCATGCAAGTGCA 843
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        318 LeuValGluCysTyrPheTrpAsnThrGlyIleIleGluProArgGlnHisAlaSerAla 337

Qy        844 AGGATAATGATGGGCAAAGTCAACGCTCTGATTACGGTGATCGATGATATTTATGATGTC 903
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        338 ArgIleMetMetGlyLysValAsnAlaLeuIleThrValIleAspAspIleTyrAspVal 357

Qy        904 TATGGCACCTTAGAAGAACTCGAACAATTCACTGACCTCATTCGAAGATGGGATATAAAC 963
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        358 TyrGlyThrLeuGluGluLeuGluGlnPheThrAspLeuIleArgArgTrpAspIleAsn 377

Qy        964 TCAATCGACCAACTTCCCGATTACATGCAACTGTGCTTTCTTGCACTCAACAACTTCGTC 1023
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        378 SerIleAspGlnLeuProAspTyrMetGlnLeuCysPheLeuAlaLeuAsnAsnPheVal 397

Qy       1024 GATGATACATCGTACGATGTTATGAAGGAGAAAGGCGTCAACGTTATACCCTACCTGCGG 1083
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        398 AspAspThrSerTyrAspValMetLysGluLysGlyValAsnValIleProTyrLeuArg 417

Qy       1084 CAATCGTGGGTTGATTTGGCGGATAAGTATATGGTAGAGGCACGGTGGTTCTACGGCGGG 1143
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        418 GlnSerTrpValAspLeuAlaAspLysTyrMetValGluAlaArgTrpPheTyrGlyGly 437

Qy       1144 CACAAACCAAGTTTGGAAGAGTATTTGGAGAACTCATGGCAGTCGATAAGTGGGCCCTGT 1203
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        438 HisLysProSerLeuGluGluTyrLeuGluAsnSerTrpGlnSerIleSerGlyProCys 457

Qy       1204 ATGTTAACGCACATATTCTTCCGAGTAACAGATTCGTTCACAAAGGAGACCGTCGACAGT 1263
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        458 MetLeuThrHisIlePhePheArgValThrAspSerPheThrLysGluThrValAspSer 477

Qy       1264 TTGTACAAATACCACGATTTAGTTCGTTGGTCATCCTTCGTTCTGCGGCTTGCTGATGAT 1323
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        478 LeuTyrLysTyrHisAspLeuValArgTrpSerSerPheValLeuArgLeuAlaAspAsp 497

Qy       1324 TTGGGAACCTCGGTGGAAGAGGTGAGCAGAGGGGATGTGCCGAAATCACTTCAGTGCTAC 1383
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        498 LeuGlyThrSerValGluGluValSerArgGlyAspValProLysSerLeuGlnCysTyr 517

Qy       1384 ATGAGTGACTACAATGCATCGGAGGCGGAGGCGCGGAAGCACGTGAAATGGCTGATAGCG 1443
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        518 MetSerAspTyrAsnAlaSerGluAlaGluAlaArgLysHisValLysTrpLeuIleAla 537

Qy       1444 GAGGTGTGGAAGAAGATGAATGCGGAGAGGGTGTCGAAGGATTCTCCATTCGGCAAAGAT 1503
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        538 GluValTrpLysLysMetAsnAlaGluArgValSerLysAspSerProPheGlyLysAsp 557

Qy       1504 TTTATAGGATGTGCAGTTGATTTAGGAAGGATGGCGCAGTTGATGTACCATAATGGAGAT 1563
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        558 PheIleGlyCysAlaValAspLeuGlyArgMetAlaGlnLeuMetTyrHisAsnGlyAsp 577

Qy       1564 GGGCACGGCACACAACACCCTATTATACATCAACAAATGACCAGAACCTTATTCGAGCCC 1623
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        578 GlyHisGlyThrGlnHisProIleIleHisGlnGlnMetThrArgThrLeuPheGluPro 597

Qy       1624 TTTGCA 1629
              ||||||
Db        598 PheAla 599

RESULT 2
H84633
probable limonene cyclase [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001
C;Accession: H84633
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, I
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter,
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: H84633
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-591 <STO>
A;Cross-references: GB:AE002093; NID:g4115381; PIDN:AAD03382.1; GSPDB:GN00139
C;Genetics:
A;Gene: At2g24210
A;Map position: 2
C;Superfamily: vetispiradiene synthase 1

Alignment Scores:
Pred. No.:              1.43e-76        Length:         591
Score:                  1147.00         Matches:        239
Percent Similarity:     64.97%          Conservative:   119
Best Local Similarity:  43.38%          Mismatches:     173
Query Match:            39.44%          Indels:         20
DB:                     2               Gaps:           10

US-09-938-956-6 (1-1632) x H84633 (1-591)

Qy          4 AGACGATCCGGAAACTACAACCCTTCTCGTTGGGATGTCAACTTCATCCAATCGCTTCTC 63
              |||||||||   ||||||   |||||||||||||||     :::    :::    |||:::
Db         46 ArgArgSerAlaAsnTyrGlnProSerArgTrpAspHisHisHisLeuLeuSerValGlu 65

Qy         64 AGTGACTATAAGGAGGACAAACACGTGATTAGGGCTTCTGAGCTGGTC------ACTTTG 117
              :::   :::   :::|||||   |||   |||      :::|||:::          :::
Db         66 AsnLysPheAlaLysAspLysArgVal---ArgGluArgAspLeuLeuLysGluLysVal 84

Qy        118 GTGAAGATGGAACTGGAGAAAGAAACGGATCAAATTCGACAACTTGAGTTGATCGATGAC 177
                 ||||||      :::::::::            :::   |||||||||   |||||||||
Db         85 ArgLysMetLeuAsnAspGluGlnLysThrTyrLeuAspGlnLeuGluPheIleAspAsp 104

Qy        178 TTGCAGAGGATGGGGCTGTCCGATCATTTCCAAAATGAGTTCAAAGAAATCTTGTCCTCT 237
```

```
              |||||::::::|||:::|||   ||||||:::   |||         ||||||:::|||
Db      105 LeuGlnLysLeuGlyValSerTyrHisPheGluAlaGluIleAspAsnIleLeuThrSer 124

Qy      238 ATATATCTCGACCATCACTATTACAAGAACCCTTTTCCAAAAGAAGAAAGGGATCTCTAC 297
                               |||  |||:::               :::|||  ||||||:::
Db      125 ---------------SerTyrLysLysAspArgThrAsnIleGlnGluSerAspLeuHis 139

Qy      298 TCCACATCTCTTGCATTTAGGCTCCTCAGAGAACATGGTTTTCAAGTCGCACAAGAGGTA 357
            :::|||:::|||   |||||||||   |||:::||||||||||   |||::::::::::|||
Db      140 AlaThrAlaLeuGluPheArgLeuPheArgGlnHisGlyPheAsnValSerGluAspVal 159

Qy      358 TTCGATAGTTTCAAGAACGAGGAGGGTGAGTTCAAAGAAAGCCTTAGCGACGACACCAGA 417
            ||||||   |||               |||:::|||                  ||||||
Db      160 PheAspValPheMetGluAsnCysGlyLysPheAspArg---------AspAspIleTyr 176

Qy      418 GGATTGTTGCAACTGTATGAAGCTTCCTTTCTGTTGACGGAAGGCGAAACCACGCTCGAG 477
            ||||||:::   |||||||||||||||:::|||   |||:::   :::      |||:::
Db      177 GlyLeuIleSerLeuTyrGluAlaSerTyrLeuSerThrLysLeuAspLysAsnLeuGln 196

Qy      478 ---TCAGCGAGGGAATTCGCCACCAAA------------TTTTTGGAGGAAAAAGTGAAC 522
                  |||   |||||||||:::               |||::::::         |||
Db      197 IlePheIleArgProPheAlaThrGlnGlnLeuArgAspPheValAspThrHisSerAsn 216

Qy      523 GAGGGTGGTGTTGATGGCGACCTTTTAACAAGAATCGCATATTCTTTGGACATCCCTCTT 582
            |||               |||::::::     :::      :::||||||:::|||
Db      217 GluAspPheGlySerCysAspMetValGluIleValValGlnAlaLeuAspMetProTyr 236

Qy      583 CATTGGAGGATTAAAAGGCCAAATGCACCTGTGTGGATCGAATGGTATAGGAAGAGGCCC 642
            :::|||:::::::::|||      :::         :::|||:::   |||   ||||||
Db      237 TyrTrpGlnMetArgArgLeuSerThrArgTrpTyrIleAspValTyrGlyLysArgGln 256

Qy      643 GAC---ATGAATCCAGTAGTGTTGGAGCTTGCCATACTCGACTTAAATATTGTTCAAGCA 699
            :::         |||   |||||||:::|||   |||   :::|||   ||||||||||||
Db      257 AsnTyrLysAsnLeuValValValGluPheAlaLysIleAspPheAsnIleValGlnAla 276

Qy      700 CAATTTCAAGAAGAGCTCAAAGAATCCTTCAGGTGGTGGAGAAATACTGGGTTTGTTGAG 759
               ||||||||||||||                  ||||||      ||||||      :::
Db      277 IleHisGlnGluGluLeuLysAsnValSerSerTrpTrpMetGluThrGlyLeuGlyLys 296

Qy      760 AAGCTGCCCTTCGCAAGGGATAGACTGGTGGAATGCTACTTTTGGAATACTGGGATCATC 819
            :::|||   |||||||||||||||:::||||||   |||||||||      |||   |||
Db      297 GlnLeuTyrPheAlaArgAspArgIleValGluAsnTyrPheTrpThrIleGlyGlnIle 316

Qy      820 GAGCCACGTCAGCATGCAAGTGCAAGGATAATGATGGGCAAAGTCAACGCTCTGATTACG 879
            :::      |||:::            |||      |||   |||:::||||||||||:::|||
Db      317 GlnGluProGlnTyrGlyTyrValArgGlnThrMetThrLysIleAsnAlaLeuLeuThr 336

Qy      880 GTGATCGATGATATTTATGATGTCTATGGCACCTTAGAAGAACTCGAACAATTCACTGAC 939
               ||||||||||||||||||||:::||||||||||||||||||||:::   ||||||
Db      337 ThrIleAspAspIleTyrAspIleTyrGlyThrLeuGluGluLeuGlnLeuPheThrVal 356

Qy      940 CTCATTCGAAGATGGGATATAAACTCAATCGACCAACTTCCCGATTACATGCAACTGTGC 999
                        ||||||||||||   :::|||:::||||||:::||||||:::||||||
Db      357 AlaPheGluAsnTrpAspIleAsnArgLeuAspGluLeuProGluTyrMetArgLeuCys 376

Qy     1000 TTTCTTGCACTCAACAACTTCGTCGATGATACATCGTACGATGTTATGAAGGAGAAAGGC 1059
            ||||||   :::   |||   |||:::      :::   ::::::::::::   |||
Db      377 PheLeuValIleTyrAsnGluValAsnSerIleAlaCysGluIleLeuArgThrLysAsn 396

Qy     1060 GTCAACGTTATACCCTACCTGCGGCAATCGTGGGTTGATTTGGCGGATAAGTATATGGTA 1119
            :::||||||||||||:::|||::::::||||||   |||::::::         |||:::|||
Db      397 IleAsnValIleProPheLeuLysLysSerTrpThrAspValSerLysAlaTyrLeuVal 416

Qy     1120 GAGGCACGGTGGTTCTACGGCGGGCACAAACCAAGTTTGGAAGAGTATTTGGAGAACTCA 1179
            ||||||:::|||:::         |||||||||||||||||||:::::::|||:::
Db      417 GluAlaLysTrpTyrLysSerGlyHisLysProAsnLeuGluGluTyrMetGlnAsnAla 436

Qy     1180 TGGCAGTCGATAAGTGGGCCCTGTATGTTAACGCACATATTCTTCCGAGTAACAGATTCG 1239
                  ||||||||||   |||   :::         |||   :::         :::|||
Db      437 ArgIleSerIleSerSerProThrIlePheValHisPheTyrCysValPheSerAspGln 456

Qy     1240 TTCACAAAGGAGACCGTCGACAGTTTGTACAAATAC---CACGATTTAGTTCGTTGGTCA 1296
               :::   :::   :::::::::|||   ::::::      ::::::||||||   |||
Db      457 LeuSerIleGlnValLeuGluThrLeuSerGlnHisGlnGlnAsnValValArgCysSer 476

Qy     1297 TCCTTCGTTCTGCGGCTTGCTGATGATTTGGGAACCTCGGTGGAAGAGGTGAGCAGAGGG 1356
            |||   |||   |||||||||:::||||||   ||||||      :::|||::::::||||||
Db      477 SerSerValPheArgLeuAlaAsnAspLeuValThrSerProAspGluLeuAlaArgGly 496

Qy     1357 GATGTGCCGAAATCACTTCAGTGCTACATGAGTGACTACAATGCATCGGAGGCGGAGGCG 1416
            ||||||   ||||||:::|||||||||||||||:::         |||||||||   :::|||
Db      497 AspValCysLysSerIleGlnCysTyrMetSerGluThrGlyAlaSerGluAspLysAla 516

Qy     1417 CGGAAGCACGTGAAATGGCTGATAGCGGAGGTGTGGAAGAAGATGAATGCGGAGAGGGTG 1476
            |||   ||||||:::   :::|||   ::::::|||   :::||||||   |||::::::
Db      517 ArgSerHisValArgGlnMetIleAsnAspLeuTrpAspGluMetAsnTyrGluLysMet 536

Qy     1477 TCGAAGGATTCTCCA---TTCGGCAAAGATTTTATAGGATGTGCAGTTGATTTAGGAAGG 1533
            :::         |||         ||||||:::         ::::::|||   |||
Db      537 AlaHisSerSerSerIleLeuHisHisAspPheMetGluThrValIleAsnLeuAlaArg 556

Qy     1534 ATGGCGCAGTTGATGTACCATAATGGAGATGGGCACGGCACA---CAACACCCTATTATA 1590
            |||:::|||   ||||||      ||||||||||||||||:::   :::         |||
Db      557 MetSerGlnCysMetTyrGlnTyrGlyAspGlyHisGlySerProGluLysAlaLysIle 576

Qy     1591 CATCAACAAATGACCAGAACCTTATTCGAGCCC 1623
                  ::::::         ||||||   |||
Db      577 ValAspArgValMetSerLeuLeuPheAsnPro 587
```

RESULT 3
F71434
probable limonene cyclase - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
A;Variety: columbia
C;Date: 03-Aug-1998 #sequence_revision 03-Aug-1998 #text_change 20-Jun-2000
C;Accession: F71434
R;Bevan, M.; Bancroft, I.; Bent, E.; Love, K.; Goodman, H.; Dean, C.; Bergkamp, R.; Dirl
 P.; Wedler, H.; Wedler, E.; Wambutt, R.; Weitzenegger, T.; Pohl, T.M.; Terryn, N.; Gie
avanagh, T.; Hempel, S.; Kotter, P.; Entian, K.D.; Rieger, M.; Schaeffer, M.; Funk, B.
Nature 391, 485-488, 1998
A;Authors: Mueller-Auer, S.; Silvey, M.; James, R.; Montfort, A.; Pons, A.; Puigdomenec
erhoft, A.; Moores, T.; Jones, J.D.G.; Eneva, T.; Palme, K.; Benes, V.; Rechman, S.; An
C.; Chalwatzis, N.
A;Title: Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of Arabidopsis tha
A;Reference number: A71400; MUID:98121113; PMID:9461215
A;Accession: F71434
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-600 <BEV>
A;Cross-references: GB:Z97341; NID:g2244991; PID:g2245028
C;Genetics:
A;Map position: 4COP9-4G3845
C;Superfamily: vetispiradiene synthase 1

```
Alignment Scores:
Pred. No.:              5.8e-72      Length:         600
Score:                  1084.50      Matches:        218
Percent Similarity:     63.34%       Conservative:   112
Best Local Similarity:  41.84%       Mismatches:     158
Query Match:            37.29%       Indels:         33
DB:                     2            Gaps:           8

US-09-938-956-6 (1-1632) x F71434 (1-600)

Qy      103 GAGCTGGTCACTTTGGTGAAGATGGAACTGGAGAAAGAAACGGATCAAATTCGACAACTT 162
            |||:::   |||      |||      |||   ||||||                  :::|||
Db       96 GluValGluThrAlaGluLysAlaIleLeuPheLysGlu---------------GluLeu 110

Qy      163 GAGTTGATCGATGACTTGCAGAGGATGGGGCTGTCCGATCATTTCCAAAATGAGTTCAAA 222
            |||:::||||||   |||||||||:::|||:::|||   |||:::::::::|||
Db      111 GluMetIleAspSerLeuGlnArgLeuGlyIleSerTyrHisTyrLysHisGluIleHis 130

Qy      223 GAAATCTTGTCCTCTATATATCTCGACCATCACTATTACAAGAACCCTTTTCCAAAAGAA 282
            :::||||||      |||:::      |||   |||               :::|||
Db      131 AspIleLeuArgLysIleHis---AspGlnHis---------GlyGluIleGluArgGlu 146
```

```
Qy      283 GAAAGGGATCTCTACTCCACATCTCTTGCATTTAGGCTCCTCAGAGAACATGGTTTTCAA 342
               :::||||||:::::::||||||||     |||     |||||||||:::||||||||
Db      147 ThrGlnAspLeuHisAlaThrSerLeuGluPheIleLeuLeuArgGlnHisGlyPheAsp 166

Qy      343 GTCGCACAAGAGGTATTCGATAGTTTCAAGAACGAGGAGGGTGAGTTCAAAGAAAGCCTT 402
            |||:::|||:::     ||||||    |||     :::|||   ||||||||::::::::|||
Db      167 ValSerGlnAspAlaPheAspValPheIleSerGluThrGlyGluPheArgLysThrLeu 186

Qy      403 AGCGACGACACCAGAGGATTGTTGCAACTGTATGAAGCTTCCTTTCTGTTGACGGAAGGC 462
                  |||     :::|||||||||    |||||||||||||||:::              :::
Db      187 HisSerAspIleLysGlyLeuLeuSerLeuTyrGluAlaSerTyrPheSerMetAspSer 206

Qy      463 GAAACCACGCTCGAGTCAGCGAGGGAATTCGCCACCAAATTTTTGGAGGAAAAAGTGAAC 522
            |||      |||:::         |||     :::|||   |||   |||   |||   |||
Db      207 GluPheLysLeuLysGluThrArgIleTyrAlaAsnLysArgLeuSerGluPheValAla 226

Qy      523 GAGGGTGGTGTT------------------GATGGCGACCTTTTAACAAGAATCGCATAT 564
            |||                              :::      :::|||      :::
Db      227 GluSerSerLysThrIleCysArgGluAspGluThrTyrIleLeuGluMetValLysArg 246

Qy      565 TCTTTGGACATCCCTCTTCATTGGAGGATTAAAAGGCCAAATGCACCTGTGTGGATCGAA 624
            :::|||:::    |||     ||||||   |||:::|||      |||        :::|||
Db      247 AlaLeuGluThrProTyrHisTrpSerIleArgArgLeuGluAlaArgTrpTyrIleAsn 266

Qy      625 TGGTATAGGAAGAGGCCCGACATGAATCCAGTAGTGTTGGAGCTTGCCATACTCGACTTA 684
               |||   |||:::   :::|||||||||::::::||||||    |||     :::|||
Db      267 ValTyrGluLysLysHisGluMetAsnProLeuLeuLeuGluPheAlaAlaIleAspPhe 286

Qy      685 AATATTGTTCAAGCACAATTTCAAGAAGAGCTCAAAGAATCCTTCAGGTGGTGGAGAAAT 744
            |||:::::::||||||         ||||||||||||||                     :::
Db      287 AsnMetLeuGlnAlaAsnHisGlnGluGluLeuLys------------LeuIleSerSer 302

Qy      745 ACTGGGTTTGTTGAGAAGCTGCCCTTCGCAAGGGATAGACTGGTGGAATGCTACTTTTGG 804
            ||||||     ::::::::::|||   |||     |||||||||:::     |||     |||||||||
Db      303 ThrGlyLeuMetLysGlnLeuAspPheValArgAspArgIleThrGluSerTyrPheTrp 322

Qy      805 AATACTGGGATCATCGAGCCACGTCAGCATGCAAGTGCAAGGATAATGATGGGCAAAGTC 864
                  ||||||              :::             |||    ::::::    |||:::
Db      323 ThrIleGlyIlePheTyrGluProGluPheLysTyrCysArgLysIleLeuThrLysIle 342

Qy      865 AACGCTCTGATTACGGTGATCGATGATATTTATGATGTCTATGGCACCTTAGAAGAACTC 924
                  ||||||     ::::::||||||||||||||||:::|||||||||||||||||||||
Db      343 PheMetLeuIleValIleMetAspAspIleTyrAspIleTyrGlyThrLeuGluGluLeu 362

Qy      925 GAACAATTCACTGACCTCATTCGAAGATGGGATATAAACTCAATCGACCAACTTCCCGAT 984
            |||     ||||||:::::::::        :::|||||||:::|||     :::::::::||||||:::
Db      363 GluLeuPheThrAsnValValGluLysTrpAspValAsnHisValGluArgLeuProAsn 382

Qy      985 TACATGCAACTGTGCTTTCTTGCACTCAACAACTTCGTCGATGATACATCGTACGATGTT 1044
            ||||||::::::|||||||||     |||     |||     ::::::         |||||||||
Db      383 TyrMetArgMetCysPheLeuPheLeuTyrAsnGluIleAsnGlnIleGlyTyrAspVal 402

Qy     1045 ATGAAGGAGAAAGGCGTCAACGTTATACCCTACCTGCGGCAATCGTGGGTTGATTTGGCG 1104
            :::::::::||||||:::||||||||||||||||||:::|||   |||     ||||||
Db      403 LeuArgAspLysGlyLeuAsnValIleProTyrLeuLysGlnValTrpThrAspLeuPhe 422

Qy     1105 GATAAGTATATGGTAGAGGCACGGTGGTTCTACGGCGGGCACAAACCAAGTTTGGAAGAG 1164
              ::::::      |||::::::|||:::               |||||||||||||||     ||||||
Db      423 LysThrPheLeuThrGluSerLysTrpTyrLysThrGlyHisLysProSerPheGluGlu 442

Qy     1165 TATTTGGAGAACTCATGGCAGTCGATAAGTGGGCCCTGTATGTTAACGCACATATTCTTC 1224
            |||::::::|||            |||   |||   |||     :::|||    |||:::|||
Db      443 TyrMetGlnAsnGlyValIleSerSerSerValProThrIleLeuLeuHisLeuPheSer 462

Qy     1225 CGAGTAACAGATTCGTTCACAAAGGAGACCGTCGACAGTTTGTACAAATACCACGATTTA 1284
              ::::::|||      :::     :::|||:::                |||   |||   :::
Db      463 ValLeuSerAspHisIleSerAspGlnThrLeuThrAspAspSerLysAsnHisSerVal 482

Qy     1285 GTTCGTTGGTCATCCTTCGTTCTGCGGCTTGCTGATGATTTGGGAACCTCGGTGGAAGAG 1344
            ||||||      :::     :::||||||||||||:::||||||     ||||||    ||||||
Db      483 ValArgSerCysAlaThrIleLeuArgLeuAlaAsnAspLeuAlaThrSerThrGluGlu 502

Qy     1345 GTGAGCAGAGGGGATGTGCCGAAATCACTTCAGTGCTACATGAGTGACTACAATGCATCG 1404
              ::::::|||||||||    |||||||||:::||||||||||||      :::        ||||||
Db      503 MetAlaArgGlyAspSerProLysSerValGlnCysTyrMetTyrGluThrArgAlaSer 522

Qy     1405 GAGGCGGAGGCGCGGAAGCACGTGAAATGGCTGATAGCGGAGGTGTGGAAGAAGATGAAT 1464
             |||     |||||||||:::|||::::::     :::|||::::::     |||       :::|||
Db      523·GluGluGluAlaArgArgHisMetGlnSerMetIleSerAspSerTrpAspIleIleAsn 542

Qy     1465 GCGGAG---AGGGTGTCGAAGGATTCTCCATTCGGCAAAGATTTTATAGGATGTGCAGTT 1521
            ::::::    :::    :::             |||           :::     |||:::     |||
Db      543 SerAspLeuLysThrAlaHisThrSerSerLeuProArgGlyPheLeuAlaAlaAlaAla 562

Qy     1522 GATTTAGGAAGGATGGCGCAGTTGATGTACCATAATGGAGATGGGCACGGCACACAACAC 1581
             :::|||   |||:::     |||      :::|||     :::|||||||||||||||:::
Db      563 AsnLeuAsnArgValValGlnCysIleTyrArgHisGlyAspGlyHisGlySer------ 580

Qy     1582 CCTATTATACATCAACAAATGACCAGAACC--------------------TTATTCGAG 1620
                        :::    |||:::|||                        ||||||
Db      581 ------------ProGluLysThrLysThrValAspTyrIleGlnSerValLeuPheAsn 596

Qy     1621 CCC 1623
            |||
Db      597 Pro 597
```

RESULT 4
G71434
probable limonene cyclase - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
A;Variety: columbia
C;Date: 03-Aug-1998 #sequence_revision 03-Aug-1998 #text_change 05-Dec-1998
C;Accession: G71434
R;Bevan, M.; Bancroft, I.; Bent, E.; Love, K.; Goodman, H.; Dean, C.; Bergkamp, R.; Dir
P.; Wedler, H.; Wedler, E.; Wambutt, R.; Weitzenegger, T.; Pohl, T.M.; Terryn, N.; Gie
avanagh, T.; Hempel, S.; Kotter, P.; Entian, K.D.; Rieger, M.; Schaeffer, M.; Funk, B.
Nature 391, 485-488, 1998
A;Authors: Mueller-Auer, S.; Silvey, M.; James, R.; Montfort, A.; Pons, A.; Puigdomenec
erhoft, A.; Moores, T.; Jones, J.D.G.; Eneva, T.; Palme, K.; Benes, V.; Rechman, S.; An
C.; Chalwatzis, N.
A;Title: Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of Arabidopsis tha
A;Reference number: A71400; MUID:98121113; PMID:9461215
A;Accession: G71434
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-1024 <BEV>
A;Cross-references: GB:Z97341; NID:g2244991; PID:e327011; PID:g2245029
C;Genetics:
A;Map position: 4COP9-4G3845

```
Alignment Scores:
Pred. No.:               1.5e-68      Length:         1024
Score:                   1039.00      Matches:        224
Percent Similarity:      62.57%       Conservative:   102
Best Local Similarity:   42.99%       Mismatches:     151
Query Match:             35.73%       Indels:         44
DB:                      2            Gaps:           14
```

US-09-938-956-6 (1-1632) x G71434 (1-1024)

```
Qy       73 AAGGAGGACAAACACGTGATTAGGGCTTCTGAGCTGGTCACTTTGGTGAAGATGGAACTG 132
            ||||||||||||                              |||     |||||||||:::|||     |||:::
Db        3 LysGluAspAsn---------------ValGluArgValThrLeuLeuLysGlnGluVal 17

Qy      133 GAGAAA---------GAAACGGAT---CAAATTCGACAACTTGAGTTGATCGATGACTTG 180
               |||               ||||||:::         :::     ||||||||||||||||||     |||
Db       18 SerLysMetLeuAsnGluThrGluGlyLeuLeuGluGlnLeuGluLeuIleAspThrLeu 37

Qy      181 CAGAGGATGGGGCTGTCCGATCATTTCCAAAATGAGTTCAAAGAAATCTTGTCCTCTATA 240
            ||||||:::|||:::|||    ||||||:::     |||     |||:::     |||:::::::::
Db       38 GlnArgLeuGlyValSerTyrHisPheGluGlnGluIleLysLysThrLeuThrAsnVal 57

Qy      241 TATCTCGACCAT---CACTATTACAAGAACCCTTTTCCAAAAGAAGAA---AGGATCTC 294
             ::::::   :::         :::||||||             :::                ||||||
Db       58 HisValLysAsnValArgAlaHisLysAsnArgIleAspArgAsnArgTrpGlyAspLeu 77
```

```
Qy    295 TACTCCACATCTCTTGCATTTAGGCTCCTCAGAGAACATGGTTTTCAAGTCGCACAAGAG 354
Db     78 TyrAlaThrAlaLeuGluPheArgLeuLeuArgGlnHis------------------Asp 91

Qy    355 GTATTCGATAGTTTCAAGAACGAGGAGGGTGAGTTCAAAGAAAGCCTTAGCGAC---GAC 411
Db     92 ValPheAsp------------------GlyAsnIleGlyValAspLeuAspAspLysAsp 105

Qy    412 ACCAGAGGATTGTTGCAACTGTATGAAGCTTCCTTTCTGTTGACGGAAGGCGAAACCACG 471
Db    106 IleLysGlyIleLeuSerLeuTyrGluAlaSerTyrLeuSerThrArgIleAspThrLys 125

Qy    472 CTCGAGTCAGCGAGGGAATTCGCCACCAAATTTTTGGAGGAA------AAAGTGAACGAG 525
Db    126 LeuLysGluSerIleTyrTyrThrThrLysArgLeuArgLysPheValGluValAsnLys 145

Qy    526 GGTGGTGTTGATGGCGACCTTTTAACAAGA---ATCGCATATTCTTTGGACATCCCTCTT 582
Db    146 AsnGluThrLysSerTyrThrLeuArgArgMetValIleHisAlaLeuGluMetProTyr 165

Qy    583 CATTGGAGGATTAAAAGGCCAAATGCACCTGTGTGGATCGAATGGTATAGGAAGAGGCCC 642
Db    166 HisArgArgValGlyArgLeuGluAlaArgTrpTyrIleGluValTyrGlyGluArgHis 185

Qy    643 GACATGAATCCAGTAGTGTTGGAGCTTGCCATACTCGACTTAAATATTGTTCAAGCACAA 702
Db    186 AspMetAsnProIleLeuLeuGluLeuAlaLysLeuAspPheAsnPheValGlnAlaIle 205

Qy    703 TTTCAAGAAGAGCTCAAAGAATCCTTCAGGTGGTGGAGAAATACTGGGTTTGTTGAGAAG 762
Db    206 HisGlnAspGluLeuLysSerLeuSerSerTrpTrpSerLysThrGlyLeuThrLysHis 225

Qy    763 CTGCCCTTCGCAAGGGATAGACTGGTGGAATGCTACTTTTGGAATACTGGGATCATCGAG 822
Db    226 LeuAspPheValArgAspArgIleThrGluGlyTyrPheSerSerValGlyValMetTyr 245

Qy    823 CCACGTCAGCATGCAAGTGCAAGGATAATGATGGGCAAAGTCAACGCTCTGATTACGGTG 882
Db    246 GluProGluPheAlaTyrHisArgGlnMetLeuThrLysValPheMetLeuIleThrThr 265

Qy    883 ATCGATGATATTTATGATGTCTATGGCACCTTAGAAGAACTCGAACAATTCACTGACCTC 942
Db    266 IleAspAspIleTyrAspIleTyrGlyThrLeuGluGluLeuGlnLeuPheThrThrIle 285

Qy    943 ATTCGAAGATGGGATATAAACTCAATCGACCAACTTCCCGATTACATGCAACTGTGCTTT 1002
Db    286 ValGluLysTrpAspValAsnArgLeuGluGluLeuProAsnTyrMetLysLeuCysPhe 305

Qy   1003 CTTGCACTCAACAACTTCGTCGATGATACATCGTACGATGTTATGAAGGAGAAAGGCGTC 1062
Db    306 LeuCysLeuValAsnGluIleAsnGlnIleGlyTyrPheValLeuArgAspLysGlyPhe 325

Qy   1063 AACGTTATACCCTACCTGCGGCAATCGTGGGTTGATTTGGCGGATAAGTATATGGTAGAG 1122
Db    326 AsnValIleProTyrLeuLysGluSerTrpAlaAspMetCysThrThrPheLeuLysGlu 345

Qy   1123 GCACGGTGGTTCTACGGCGGGCACAAACCAAGTTTGGAAGAGTATTTGGAGAACTCATGG 1182
Db    346 AlaLysTrpTyrLysSerGlyTyrLysProAsnPheGluGluTyrMetGlnAsnGlyTrp 365

Qy   1183 CAGTCGATAAGTGGGCCCTGTATGTTAACGCACATATTCTTCCGAGTAACAGATTCGTTC 1242
Db    366 IleSerSerSerValProThrIleLeuLeuHisLeuPheCysLeuLeuSerAsp------ 383

Qy   1243 ACAAAGGAGACCGTCGACAGTTTGTACAAATAC---CACGATTTAGTTCGTTGGTCATCC 1299
Db    384 ------GlnThrLeuAspIleLeuGlySerTyrAsnHisSerValValArgSerSerAla 401

Qy   1300 TTCGTTCTGCGGCTTGCTGATGATTTGGGAACCTCG------------------------ 1335
Db    402 ThrIleLeuArgLeuAlaAsnAspLeuAlaThrSerSerValSerHisGlyPheThrThr 421

Qy   1336 ------GTGGAAGAGGTGAGCAGAGGGGATGTGCCGAAATCACTTCAGTGCTACATGAGT 1389
Db    422 TyrAsnThrGluGluLeuAlaArgGlyAspThrMetLysSerValGlnCysHisMetHis 441

Qy   1390 GACTACAATGCATCGGAGGCGGAGGCGCGGAAGCACGTGAAATGGCTGATAGCGGAGGTG 1449
Db    442 GluThrGlyAlaSerGluAlaGluSerArgAlaTyrIleGlnGlyIleIleGlyValAla 461

Qy   1450 TGGAAGAAGATGAATGCGGAGAGGGTGTCGAAGGATTCTCCATTCGGCAAAGATTTTATA 1509
Db    462 TrpAspAspLeuAsnMetGluLys------LysSerCysArgLeuHisGlnGlyPheLeu 479

Qy   1510 GGATGTGCAGTTGATTTAGGAAGGATGGCGCAGTTGATGTACCATAATGGAGATGGGCAC 1569
Db    480 GluAlaAlaAlaAsnLeuGlyArgValAlaGlnCysValTyrGlnTyrGlyAspGlyHis 499

Qy   1570 GGC 1572
Db    500 Gly 500
```

RESULT 5
T08174
sesquiterpene cyclase (EC 2.5.1.-) - pepper
C;Species: Capsicum annuum (pepper)
C;Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000
C;Accession: T08174
R;Back, K.; Shin, D.H.; He, S.
Plant Cell Physiol. 39, 899-904, 1998
A;Title: Cloning and bacterial expression of sesquiterpene cyclase, a key branch point
A;Reference number: Z16395; MUID:99033462; PMID:9816674
A;Accession: T08174
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-559 <BAC>
A;Cross-references: EMBL:AF061285; NID:g3108342; PIDN:AAC61260.1; PID:g3108343
C;Function:
A;Pathway: the synthesis of phytoalexin capsidiol
A;Note: unduced by UV
C;Superfamily: vetispiradiene synthase 1
C;Keywords: isoprenoid biosynthesis; transferase

```
Alignment Scores:
Pred. No.:               1.18e-52      Length:          559
Score:                   822.50        Matches:         182
Percent Similarity:      52.85%        Conservative:    115
Best Local Similarity:   32.38%        Mismatches:      222
Query Match:             28.28%        Indels:          43
DB:                      2             Gaps:            9

US-09-938-956-6 (1-1632) x T08174 (1-559)

Qy      1 ATGAGACGATCCGGAAACTACAACCCTTCTCGTTGGGATGTCAACTTCATCCAATCGCTT 60
Db     19 IleArgProValAlaAspPheSerProSerLeuTrpGlyAspArgPheLeuSerPheSer 38

Qy     61 CTCAGTGACTATAAGGAGGACAAACACGTGATTAGGGCTTCTGAGCTGGTCACTTTGGTG 120
Db     39 IleAspAsnGlnValGluThrLysTyrAla------------------------------ 48

Qy    121 AAGATGGAACTGGAGAAAGAAACGGATCAAATTCGACAA--------------------- 159
Db     49 ---GlnGluIleGluProLeuLysGluGlnThrArgSerMetLeuLeuAlaSerGlyArg 67

Qy    160 ---------------CTTGAGTTGATCGATGACTTGCAGAGGATGGGGCTGTCCGATCAT 204
Db     68 LysLeuSerGluThrLeuAsnLeuIleAspValIleGluArgLeuGlyIleAlaTyrHis 87

Qy    205 TTCCAAAATGAGTTCAAAGAAATCTTGTCCTCTATATATCTCGACCATCACTATTACAAG 264
Db     88 PheGluLysGluIleAspGluIleLeuAspArgIleTyrAsnGluAsnSerAsnPheGlu 107

Qy    265 AACCCTTTTCCAAAAGAAGAAAGGGATCTCTACTCCACATCTCTTGCATTTAGGCTCCTC 324
Db    108 GlyAspValTyrAsnGlu------AspLeuCysThrCysArgLeuGlnPheArgLeuLeu 125
```

```
Qy    325 AGAGAACATGGTTTTCAAGTCGCACAAGAGGTATTCGATAGTTTCAAGAACGAGGAGGGT 384
Db    126 ArgGlnHisGlyTyrAsnIleSerLeuLysIlePheSerLysPheLeuAspGlyAsnGly 145

Qy    385 GAGTTCAAAGAAAGCCTTAGCGACGACACCAGAGGATTGTTGCAACTGTATGAAGCTTCC 444
Db    146 ArgLeuLysGluSerLeuAlaSerAspValLeuGlyLeuLeuSerLeuTyrGluAlaSer 165

Qy    445 TTTCTGTTGACGGAAGGCGAAACCACGCTCGAGTCAGCGAGGGAATTCGCCACCAAATTT 504
Db    166 HisValArgSerHisGlyGluAspIleLeuGluAspAlaLeuAlaPheSerThrThrHis 185

Qy    505 TTGGAGGAAAAAGTGAACGAGGGTGGTGTTGATGGCGACCTTTTAACAAGAATCGCATAT 564
Db    186 LeuGluSerAlaThrProHis------LeuGluTyrProLeuLysGluGlnValArgHis 203

Qy    565 TCTTTGGACATCCCTCTTCATTGGAGGATTAAAAGGCCAAATGCACCTGTGTGGATC--- 621
Db    204 AlaLeuGluGlnSerLeuHisLysGlyIleProArgIleGluIleGlnPhePheIleSer 223

Qy    622 GAATGGTATAGGAAGAGGCCCGACATGAATCCAGTAGTGTTGGAGCTTGCCATACTCGAC 681
Db    224 SerValTyrAspLysGlnAlaIleLysAsnAspValLeuLeuArgPheAlaLysLeuAsp 243

Qy    682 TTAAATATTGTTCAAGCACAATTTCAAGAAGAGCTCAAAGAATCCTTCAGGTGGTGGAGA 741
Db    244 TyrAsnMetLeuGlnMetLeuHisLysGlnGluLeuAlaGluValSerArgTrpTrpLys 263

Qy    742 AATACTGGGTTTGTTGAGAAGCTGCCCTTCGCAAGGGATAGACTGGTGGAATGCTACTTT 801
Db    264 AspLeuAsnPheValAsnThrLeuProTyrAlaArgAspArgValValGluCysTyrPhe 283

Qy    802 TGGAATACTGGGATCATCGAGCCACGTCAGCATGCAAGTGCAAGGATAATGATGGGCAAA 861
Db    284 TrpAlaLeuGlyValTyrTyrGluProGlnTyrSerGlnAlaArgValMetLeuValLys 303

Qy    862 GTCAACGCTCTGATTACGGTGATCGATGATATTTATGATGTCTATGGCACCTTAGAAGAA 921
Db    304 ThrIleAlaMetIleSerIleValAspAspThrTyrAspAlaTyrGlyThrValAspGlu 323

Qy    922 CTCGAACAATTCACTGACCTCATTCGAAGATGGGATATAAACTCAATCGACCAACTTCCC 981
Db    324 LeuAlaIleTyrThrAspValIleGlnArgTrpAspIleLysGluIleAspSerLeuPro 343

Qy    982 GATTACATGCAACTGTGCTTTCTTGCACTCAACAACTTCGTCGATGATACATCGTACGAT 1041
Db    344 AspTyrMetLysIleSerTyrLysAlaLeuLeuAspLeuTyrLysAspTyrGluLysGlu 363

Qy   1042 GTTATGAAGGAGAAAGGCGTCAACGTTATACCCTACCTGCGGCAATCGTGGGTTGATTTG 1101
Db    364 MetSerArgAspGlyArgSerHisValValTyrTyrAlaLysGluArgLeuLysGluLeu 383

Qy   1102 GCGGATAAGTATATGGTAGAGGCACGGTGGTTCTACGGCGGGCACAAACCAAGTTTGGAA 1161
Db    384 ValLysSerTyrAsnIleGluAlaLysTrpPheIleGluGlyHisMetProProAlaSer 403

Qy   1162 GAGTATTTGGAGAACTCATGGCAGTCGATAAGTGGGCCCTGTATGTTAACGCACATATTC 1221
Db    404 GluTyrLeuArgAsnAlaPhe---------------------ValThrThrThrTyrTyr 416

Qy   1222 TTCCGAGTAACAGATTCGTTC------------------ACAAAGGAGACCGTCGACAGT 1263
Db    417 TyrLeuAlaThrThrSerTyrLeuGlyMetLysTyrAlaLysGluGlnGlnPheGluTrp 436

Qy   1264 TTGTACAAATACCACGATTTAGTTCGTTGGTCATCCTTCGTTCTGCGGCTTGCTGATGAT 1323
Db    437 LeuSerLysAsnProLysIleLeuGluGlyCysValThrIleCysArgValIleAspAsp 456

Qy   1324 TTGGGAACCTCGGTGGAAGAGGTGAGCAGAGGGGATGTGCCGAAATCACTTCAGTGCTAC 1383
Db    457 IleAlaThrTyrGluValGluLysAsnArgGlyGlnLeuSerThrGlyIleGluCysTyr 476

Qy   1384 ATGAGTGACTACAATGCATCGGAGGCGGAGGCGCGGAAGCACGTGAAATGGCTGATAGCG 1443
Db    477 MetArgAspTyrSerValSerThrLysGluAlaMetAlaLysPheGlnGluMetGlyGlu 496

Qy   1444 GAGGTGTGGAAGAAGATGAATGCGGAGAGGGTGTCGAAGGATTCTCCATTCGGCAAAGAT 1503
Db    497 SerGlyTrpLysAspIleAsn---GluGlyMetLeuArgProThrProIleProMetGlu 515

Qy   1504 TTTATAGGATGTGCAGTTGATTTAGGAAGGATGGCGCAGTTGATGTAC---CATAATGGA 1560
Db    516 PheLeuSerArgIleLeuAsnLeuAlaArgLeuValAspValThrTyrLysHisAsnGlu 535

Qy   1561 GATGGGCACGGCACACAACACCCTATTATACATCAACAAATGACCAGAACCTTATTCGAG 1620
Db    536 AspGlyTyrThrHisProGluLysValIleLysProHisIleIleAlaMetValValAsp 555

Qy   1621 CCCTTT 1626
Db    556 SerPhe 557
```

RESULT 6
A56118
vetispiradiene synthase 1 - Hyoscyamus muticus (fragment)
C;Species: Hyoscyamus muticus
C;Date: 19-Oct-1995 #sequence_revision 19-Oct-1995 #text_change 29-Sep-1999
C;Accession: A56118
R;Back, K.; Chappell, J.
J. Biol. Chem. 270, 7375-7381, 1995
A;Title: Cloning and bacterial expression of a sesquiterpene cyclase from Hyoscyamus mu
A;Reference number: A56118; MUID:95221394; PMID:7706281
A;Accession: A56118
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-520 <BAC>
A;Cross-references: GB:U20188; NID:g763422; PIDN:AAA86337.1; PID:g763423
C;Superfamily: vetispiradiene synthase 1

Alignment Scores:

Pred. No.:	6.88e-52	Length:	520
Score:	812.00	Matches:	170
Percent Similarity:	56.61%	Conservative:	104
Best Local Similarity:	35.12%	Mismatches:	196
Query Match:	27.92%	Indels:	14
DB:	2	Gaps:	7

US-09-938-956-6 (1-1632) x A56118 (1-520)

```
Qy    142 ACGGATCAAATTCGACAACTTGAGTTGATCGATGACTTGCAGAGGATGGGGCTGTCCGAT 201
Db     30 ThrThrLeuThrGluLysLeuAsnLeuIleAspIleIleGluArgLeuGlyIleAlaTyr 49

Qy    202 CATTTCCAAAATGAGTTCAAAGAAATCTTGTCCTCTATATATCTCGACCATCACTATTAC 261
Db     50 HisPheGluLysGlnIleGluAspMet---------------LeuAspHisIleTyrArg 64

Qy    262 AAGAACCCTTTT------CCAAAAGAAGAAAGGATCTCTACTCCACATCTCTTGCATTT 315
Db     65 AlaAspProTyrPheGluAlaHisGluTyrAsnAspLeuAsnThrSerSerValGlnPhe 84

Qy    316 AGGCTCCTCAGAGAACATGGTTTTCAAGTCGCACAAGAGGTATTCGATAGTTTCAAGAAC 375
Db     85 ArgLeuLeuArgGlnHisGlyTyrAsnValSerProAsnIlePheSerArgPheGlnAsp 104

Qy    376 GAGGAGGGTGAGTTCAAAGAAAGCCTTAGCGACGACACCAGAGGATTGTTGCAACTGTAT 435
Db    105 AlaAsnGlyLysPheLysGluSerLeuArgSerAspIleArgGlyLeuLeuAsnLeuTyr 124

Qy    436 GAAGCTTCCTTTCTGTTGACGGAAGGCGAAACCACGCTCGAGTCAGCGAGGGAATTCGCC 495
Db    125 GluAlaSerHisValArgThrHisLysGluAspIleLeuGluGluAlaLeuValPheSer 144

Qy    496 ACCAAATTTTTGGAGGAAAAAGTGAACGAGGGTGGTGTTGATGGCGACCTTTTAACAAGA 555
```

```
Db    145 ValGlyHisLeuGluSerAlaAlaProHis------LeuLysSerProLeuSerLysGln 162
Qy    556 ATCGCATATTCTTTGGACATCCCTCTTCATTGGAGGATTAAAAGGCCAAATGCACCTGTG 615
Db    163 ValThrHisAlaLeuGluGlnSerLeuHisLysSerIleProArgValGluIleArgTyr 182
Qy    616 TGGATCGAATGGTATAGGAAGAGGCCCGACATGAATCCAGTAGTGTTGGAGCTTGCCATA 675
Db    183 PheIleSerIleTyrGluGluGluPheLysAsnAspLeuLeuLeuArgPheAlaLys 202
Qy    676 CTCGACTTAAATATTGTTCAAGCACAATTTCAAGAAGAGCTCAAAGAATCCTTCAGGTGG 735
Db    203 LeuAspTyrAsnLeuLeuGlnMetLeuHisLysHisGluLeuSerGluValSerArgTrp 222
Qy    736 TGGAGAAATACTGGGTTTGTTGAGAAGCTGCCCTTCGCAAGGGATAGACTGGTGGAATGC 795
Db    223 TrpLysAspLeuAspPheValThrThrLeuProTyrAlaArgAspArgAlaValGluCys 242
Qy    796 TACTTTTGGAATACTGGGATCATCGAGCCACGTCAGCATGCAAGTGCAAGGATAATGATG 855
Db    243 TyrPheTrpThrMetGlyValTyrAlaGluProGlnTyrSerGlnAlaArgValMetLeu 262
Qy    856 GGCAAAGTCAACGCTCTGATTACGGTGATCGATGATATTTATGATGTCTATGGCACCTTA 915
Db    263 AlaLysThrIleAlaMetIleSerIleValAspAspThrPheAspAlaTyrGlyIleVal 282
Qy    916 GAAGAACTCGAACAATTCACTGACCTCATTCGAAGATGGGATATAAACTCAATCGACCAA 975
Db    283 LysGluLeuGluValTyrThrAspAlaIleGlnArgTrpAspIleSerGlnIleAspArg 302
Qy    976 CTTCCCGATTACATGCAACTGTGCTTTCTTGCACTCAACAACTTCGTCGATGATACATCG 1035
Db    303 LeuProGluTyrMetLysIleSerTyrLysAlaLeuLeuAspLeuTyrAspAspTyrGlu 322
Qy   1036 TACGATGTTATGAAGGAGAAAGGCGTCAACGTTATACCCTACCTGCGGCAATCGTGGGTT 1095
Db    323 LysGluLeuSerLysAspGlyArgSerAspValValHisTyrAlaLysGluArgMetLys 342
Qy   1096 GATTTGGCGGATAAGTATATGGTAGAGGCACGGTGGTTCTACGGCGGGCACAAACCAAGT 1155
Db    343 GluIleValGlyAsnTyrPheIleGluGlyLysTrpPheIleGluGlyTyrMetProSer 362
Qy   1156 TTGGAAGAGTATTTGGAGAACTCATGGCAGTCGATAAGTGGGCCCTGTATGTTAACGCAC 1215
Db    363 ValSerGluTyrLeuSerAsnAla---LeuAlaThrSerThrTyrTyrLeuLeuThrThr 381
Qy   1216 ATATTCTTCCGAGTAACAGATTCGTTCACAAAGGAGACCGTCGACAGTTTGTACAAATAC 1275
Db    382 ThrSerTyrLeuGlyMetLysSerAlaThrLysGluHisPheGluTrpLeuAlaThrAsn 401
Qy   1276 CACGATTTAGTTCGTTGGTCATCCTTCGTTCTGCGGCTTGCTGATGATTTGGGAACCTCG 1335
Db    402 ProLysIleLeuGluAlaAsnAlaThrLeuCysArgValValAspAspIleAlaThrTyr 421
Qy   1336 GTGGAAGAGGTGAGCAGAGGGGATGTGCCGAAATCACTTCAGTGCTACATGAGTGACTAC 1395
Db    422 GluValGluLysGlyArgGlyGlnIleAlaThrGlyIleGluCysTyrMetArgAspTyr 441
Qy   1396 AATGCATCGGAGGCGGAGGCGCGGAAGCACGTGAAATGGCTGATAGCGGAGGTGTGGAAG 1455
Db    442 GlyValSerThrGluValAlaMetGluLysPheGlnGluMetAlaAspIleAlaTrpLys 461
Qy   1456 AAGATGAATGCGGAGAGGGTGTCGAAGGATTCTCCATTCGGCAAAGATTTTATAGGATGT 1515
Db    462 AspValAsn---GluGluIleLeuArgProThrProValSerSerGluIleLeuThrArg 480
Qy   1516 GCAGTTGATTTAGGAAGGATGGCGCAGTTGATGTAC---CATAATGGAGATGGGCACGGC 1572
Db    481 IleLeuAsnLeuAlaArgIleIleAspValThrTyrLysHisAsnGlnAspGlyTyr--- 499
Qy   1573 ACACAACACCCT 1584
Db    500 ---ThrHisPro 502
```

RESULT 7
T03714
5-epi-aristolochene synthase - common tobacco
N;Alternate names: sesquiterpene cyclase
C;Species: Nicotiana tabacum (common tobacco)
C;Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 29-Sep-1999
C;Accession: T03714
R;Facchini, P.J.; Chappell, J.
Proc. Natl. Acad. Sci. U.S.A. 89, 11088-11092, 1992
A;Title: Gene family for an elicitor-induced sesquiterpene cyclase in tobacco.
A;Reference number: Z15024; MUID:93066390; PMID:1438319
A;Accession: T03714
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-550 <FAC>
A;Cross-references: EMBL:L04680; NID:g170342; PIDN:AAA19216.1; PID:g505588
A;Experimental source: strain NK326
C;Genetics:
A;Introns: 37/3; 126/1; 251/2; 324/2; 370/3; 452/3
C;Function:
A;Description: mediates the conversion of the isoprenoid intermediate farnesyl diphosph
C;Superfamily: vetispiradiene synthase 1

Alignment Scores:

Pred. No.:	1.07e-51	Length:	550
Score:	809.50	Matches:	179
Percent Similarity:	53.10%	Conservative:	112
Best Local Similarity:	32.66%	Mismatches:	208
Query Match:	27.84%	Indels:	49
DB:	2	Gaps:	10

US-09-938-956-6 (1-1632) x T03714 (1-550)

```
Qy      1 ATGAGACGATCCGGAAACTACAACCCTTCTCGTTGGGATGTCAACTTCATCCAATCGCTT 60
Db     14 ValArgProValAlaAspPheSerProSerLeuTrpGlyAspGlnPheLeuSerPheSer 33
Qy     61 CTCAGTGACTATAAGGAGGACAAACACGTGATTAGGGCTTCTGAGCTGGTCACTTTGGTG 120
Db     34 IleAspAsnGlnValAlaGluLysTyrIle-------------------------Tyr 44
Qy    121 AAGATGGAACTGGAGAAAGAAACGGATCAAATTCGACAA--------------------- 159
Db     45 AlaGlnGluIleGluAlaLeuLysGluGlnThrArgSerMetLeuLeuAlaThrGlyArg 64
Qy    160 ---------------CTTGAGTTGATCGATGACTTGCAGAGGATGGGGCTGTCCGATCAT 204
Db     65 LysLeuAlaAspThrLeuAsnLeuIleAspIleIleGluArgLeuGlyIleSerTyrHis 84
Qy    205 TTCCAAAATGAGTTCAAAGAAATCTTGTCCTCTATATATCTCGACCATCACTATTACAAG 264
Db     85 PheGluLysGluIleAspGluIleLeuAspGlnIleTyrAsnGlnAsn------------ 100
Qy    265 AACCCTTTTCCAAAAGAAGAAAGGGATCTCTACTCCACATCTCTTGCATTTAGGCTCCTC 324
Db    101 ------------SerAsnCysAsnAspLeuCysThrSerAlaLeuGlnPheArgLeuLeu 116
Qy    325 AGAGAACATGGTTTTCAAGTCGCACAAGAGGTATTCGATAGTTTCAAGAACGAGGAGGGT 384
Db    117 ArgGlnHisGlyPheAsnIleSerProGluIlePheSerLysPheGlnAspGluAsnGly 136
Qy    385 GAGTTCAAAGAAAGCCTTAGCGACGACACCAGAGGATTGTTGCAACTGTATGAAGCTTCC 444
Db    137 LysPheLysGluSerLeuAlaSerAspValLeuGlyLeuLeuAsnLeuTyrGluAlaSer 156
Qy    445 TTTCTGTTGACGGAAGGCGAAACCACGCTCGAGTCAGCGAGGGAATTCGCCACCAAATTT 504
Db    157 HisValArgThrHisAlaAspAspIleLeuGluAspAlaLeuAlaPheSerThrIleHis 176
Qy    505 TTGGAGGAAAAAGTGAACGAGGGTGGTGTTGATGGCGACCTTTTAACAAGAATCGCATAT 564
Db    177 LeuGluSerAlaAlaProHis------LeuLysSerProLeuArgGluGlnValThrHis 194
```

```
Qy        565 TCTTTGGACATCCCTCTTCATTGGAGGATTAAAAGGCCAAATGCACCTGTGTGGATC--- 621
Db        195 AlaLeuGluGlnCysLeuHisLysGlyValProArgValGluThrArgPhePheIleSer 214
Qy        622 GAATGGTATAGGAAGAGGCCCGACATGAATCCAGTAGTGTTGGAGCTTGCCATACTCGAC 681
Db        215 SerIleTyrAspLysGluGlnSerLysAsnAsnValLeuLeuArgPheAlaLysLeuAsp 234
Qy        682 TTAAATATTGTTCAAGCACAATTTCAAGAAGAGCTCAAAGAATCCTTCAGGTGGTGGAGA 741
Db        235 PheAsnLeuLeuGlnMetLeuHisLysGlnGluLeuAlaGlnValSerArgTrpTrpLys 254
Qy        742 AATACTGGGTTTGTTGAGAAGCTGCCCTTCGCAAGGGATAGACTGGTGGAATGCTACTTT 801
Db        255 AspLeuAspPheValThrThrLeuProTyrAlaArgAspArgValValGluCysTyrPhe 274
Qy        802 TGGAATACTGGGATCATCGAGCCACGTCAGCATGCAAGTGCAAGGATAATGATGGGCAAA 861
Db        275 TrpAlaLeuGlyValTyrPheGluProGlnTyrSerGlnAlaArgValMetLeuValLys 294
Qy        862 GTCAACGCTCTGATTACGGTGATCGATGATATTTATGATGTCTATGGCACCTTAGAAGAA 921
Db        295 ThrIleSerMetIleSerIleValAspAspThrPheAspAlaTyrGlyThrValLysGlu 314
Qy        922 CTCGAACAATTCACTGACCTCATTCGAAGATGGGATATAAACTCAATCGACCAACTTCCC 981
Db        315 LeuGluAlaTyrThrAspAlaIleGlnArgTrpAspIleAsnGluIleAspArgLeuPro 334
Qy        982 GATTACATGCAACTGTGCTTTCTTGCACTCAACAACTTCGTCGATGATACATCGTACGAT 1041
Db        335 AspTyrMetLysIleSerTyrLysAlaIleLeuAspLeuTyrLysAspTyrGluLysGlu 354
Qy       1042 GTTATGAAGGAGAAAGGCGTCAACGTTATACCCTACCTGCGGCAATCGTGGGTTGATTTG 1101
Db        355 LeuSerSerAlaGlyArgSerHisIleValCysHisAlaIleGluArgMetLysGluVal 374
Qy       1102 GCGGATAAGTATATGGTAGAGGCACGGTGGTTCTACGGCGGCACAAACCAAGTTTGGAA 1161
Db        375 ValArgAsnTyrAsnValGluSerThrTrpPheIleGluGlyTyrMetProProValSer 394
Qy       1162 GAGTATTTGGAGAACTCATGGCAGTCGATAAGTGGGCCCTGTATGTTAACGCACATATTC 1221
Db        395 GluTyrLeuSerAsnAlaLeuAlaThr--------------------ThrThrTyrTyr 407
Qy       1222 TTCCGAGTAACAGATTCGTTC-----------------ACAAAGGAGACCGTCGACAGT 1263
Db        408 TyrLeuAlaThrThrSerTyrLeuGlyMetLysSerAlaThrGluGlnAspPheGluTrp 427
Qy       1264 TTGTACAAATACCACGATTTAGTTCGTTGGTCATCCTTCGTTCTGCGGCTTGCTGATGAT 1323
Db        428 LeuSerLysAsnProLysIleLeuGluAlaSerValIleIleCysArgValIleAspAsp 447
Qy       1324 TTGGGAACCTCGGTGGAAGAGGTGAGCAGAGGGGATGTGCCGAAATCACTTCAGTGCTAC 1383
Db        448 ThrAlaThrTyrGluValGluLysSerArgGlyGlnIleAlaThrGlyIleGluCysCys 467
Qy       1384 ATGAGTGACTACAATGCATCGGAGGCGGAGGCGCGGAAGCACGTGAAATGGCTGATAGCG 1443
Db        468 MetArgAspTyrGlyIleSerThrLysGluAlaMetAlaLysPheGlnAsnMetAlaGlu 487
Qy       1444 GAGGTGTGGAAGAAGATGAATGCGGAGAGGGTGTCGAAGGATTCTCCATTCGGCAAAGAT 1503
Db        488 ThrAlaTrpLysAspIleAsn---GluGlyLeuLeuArgProThrProValSerThrGlu 506
Qy       1504 TTTATAGGATGTGCAGTTGATTTAGGAAGGATGGCGCAGTTGATGTAC---CATAATGGA 1560
Db        507 PheLeuThrProIleLeuAsnLeuAlaArgIleValGluValThrTyrIleHisAsnLeu 526
Qy       1561 GATGGGCACGGCACACAACACCCT 1584
Db        527 AspGlyTyr------ThrHisPro 532
```

RESULT 8
S68366
(+)-delta-cadinene synthase isozyme XC14 - Gossypium arboreum
C;Species: Gossypium arboreum
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 29-Sep-1999
C;Accession: S68366
R;Chen, X.Y.; Chen, Y.; Heinstein, P.; Davisson, V.J.
Arch. Biochem. Biophys. 324, 255-266, 1995
A;Title: Cloning, expression, and characterization of (+)-delta-cadinene synthase: a ca
A;Reference number: S68365; MUID:96132653; PMID:8554317
A;Accession: S68366
A;Molecule type: mRNA
A;Residues: 1-554 <CHE>
A;Cross-references: EMBL:U23205; NID:g1045313; PIDN:AAA93065.1; PID:g1045314
A;Experimental source: cultivar Nanking
C;Superfamily: vetispiradiene synthase 1
C;Keywords: phytoalexin biosynthesis

Alignment Scores:

Pred. No.:	1.06e-50	Length:	554
Score:	796.00	Matches:	185
Percent Similarity:	52.07%	Conservative:	104
Best Local Similarity:	33.33%	Mismatches:	232
Query Match:	27.37%	Indels:	34
DB:	2	Gaps:	7

US-09-938-956-6 (1-1632) x S68366 (1-554)

```
Qy          1 ATGAGACGATCCGGAAACTACAACCCTTCTCGTTGG----------------------- 36
Db         21 MetArgProLysAlaAspPheGlnProSerIleTrpGlyAspLeuPheLeuAsnCysPro 40
Qy         37 GATGTCAACTTCATCCAATCGCTTCTCAGTGACTATAAGGAGGACAAACACGTGATTAGG 96
Db         41 AspLysAsnIleAspAlaGluThrGluLysArgHisGlnGlnLeuLysGluGluValArg 60
Qy         97 GCTTCTGAGCTGGTCACTTTGGTGAAGATGGAACTGGAGAAAGAAACGGATCAAATTCGA 156
Db         61 -----------------------LysMetIleValAlaProMetAlaAsnSerThrGln 72
Qy        157 CAACTTGAGTTGATCGATGACTTGCAGAGGATGGGGCTGTCCGATCATTTCCAAAATGAG 216
Db         73 LysLeuAlaPheIleAspSerValGlnArgLeuGlyValSerTyrHisPheThrLysGlu 92
Qy        217 TTCAAAGAAATCTTGTCCTCTATATATCTCGACCATCACTATTACAAGAACCCTTTTCCA 276
Db         93 IleGluAspGluLeuGluAsnIleTyrHisAsnAsnAsn-------------------- 105
Qy        277 AAAGAAGAAAGGGATCTCTACTCCACATCTCTTGCATTTAGGCTCCTCAGAGAACATGGT 336
Db        106 AspAlaGluAsnAspLeuTyrThrThrSerLeuArgPheArgLeuLeuArgGluHisGly 125
Qy        337 TTTCAAGTCGCACAAGAGGTATTCGATAGTTTCAAGAACGAGGAGGGTGAGTTCAAAGAA 396
Db        126 PheAsnValSerCysAspValPheAsnLysPheLysAspGluGlnGlyAsnPheLysSer 145
Qy        397 AGCCTTAGCGACGACACCAGAGGATTGTTGCAACTGTATGAAGCTTCCTTTCTGTTGACG 456
Db        146 SerValThrSerAspValArgGlyLeuLeuGluLeuTyrGlnAlaSerTyrLeuArgVal 165
Qy        457 GAAGGCGAAACCACGCTCGAGTCAGCGAGGGAATTCGCCACCAAATTTTTGGAGGAAAAA 516
Db        166 HisGlyGluAspIleLeuAspGluAlaIleSerPheThrThrAsnHisLeuSerLeuAla 185
Qy        517 GTGAACGAGGGTGGTGTTGATGGCGACCTTTTAACAAGAATCGCATATTCTTTGGACATC 576
Db        186 Val------AlaSerLeuAspTyrProLeuSerGluGluValSerHisAlaLeuLysGln 203
Qy        577 CCTCTTCATTGGAGGATTAAAAGGCCAAATGCACCTGTGTGGATCGAATGGTATAGGAAG 636
Db        204 SerIleArgArgGlyLeuProArgValGluAlaArgHisTyrLeuSerValTyrGlnAsp 223
Qy        637 AGGCCCGACATGAATCCAGTAGTGTTGGAGCTTGCCATACTCGACTTAAATATTGTTCAA 696
```

```
Db     224 IleGluSerHisAsnLysValLeuLeuGluPheAlaLysIleAspPheAsnMetValGln 243
Qy     697 GCACAATTTCAAGAAGAGCTCAAAGAATCCTTCAGGTGGTGGAGAAATACTGGGTTTGTT 756
Db     244 LeuLeuHisArgLysGluLeuSerGluIleSerArgTrpTrpLysAspLeuAspPheGln 263
Qy     757 GAGAAGCTGCCCTTCGCAAGGGATAGACTGGTGGAATGCTACTTTTGGAATACTGGGATC 816
Db     264 ArgLysLeuProTyrAlaArgAspArgValValGluGlyTyrPheTrpIleSerGlyVal 283
Qy     817 ATCGAGCCACGTCAGCATGCAAGTGCAAGGATAATGATGGGCAAAGTCAACGCTCTGATT 876
Db     284 TyrPheGluProGlnTyrSerLeuGlyArgLysMetLeuThrLysValIleAlaMetAla 303
Qy     877 ACGGTGATCGATGATATTTATGATGTCTATGGCACCTTAGAAGAACTCGAACAATTCACT 936
Db     304 SerIleValAspAspThrTyrAspSerTyrAlaThrTyrGluGluLeuIleProTyrThr 323
Qy     937 GACCTCATTCGAAGATGGGATATAAACTCAATCGACCAACTTCCCGATTACATGCAACTG 996
Db     324 LysAlaIleGluArgTrpAspIleLysCysIleAspGluLeuProGluTyrMetLysPro 343
Qy     997 TGCTTTCTTGCACTCAACAACTTCGTCGATGATACATCGTACGATGTTATGAAGGAGAAA 1056
Db     344 SerTyrLysAlaLeuLeuAspValTyrGluGluMetGluGlnLeuValAlaLysHisGly 363
Qy    1057 GGCGTCAACGTTATACCCTACCTGCGGCAATCGTGGGTTGATTTGGCGGATAAGTATATG 1116
Db     364 ArgGlnTyrArgValGluTyrAlaLysAsnAlaMetIleArgLeuAlaGlnSerTyrLeu 383
Qy    1117 GTAGAGGCACGGTGGTTCTACGGCGGGCACAAACCAAGTTTGGAAGAGTATTTGGAGAAC 1176
Db     384 ValGluAlaArgTrpThrLeuGlnAsnTyrLysProSerPheGluGluPheLysAlaAsn 403
Qy    1177 TCATGGCAGTCGATAAGTGGGCCCTGTATGTTAACGCACATATTCTTCCGAGTAACAGAT 1236
Db     404 AlaLeuProThrCysGlyTyrAlaMetLeuAlaIleThrSerPheValGlyMetGlyAsp 423
Qy    1237 TCGTTCACAAAGGAGACCGTCGACAGTTTGTACAAATACCACGATTTAGTTCGTTGGTCA 1296
Db     424 IleValThrProGluThrPheLysTrpAlaAlaAsnAspProLysIleIleGlnAlaSer 443
Qy    1297 TCCTTCGTTCTGCGGCTTGCTGATGATTTGGGAACCTCGGTGGAAGAGGTGAGCAGAGGG 1356
Db     444 ThrIleIleCysArgPheMetAspAspValAlaGluHisLysPheLysHisArgArgGlu 463
Qy    1357 GATGTGCCGAAATCACTTCAGTGCTACATGAGTGACTACAATGCATCGGAGGCGGAGGCG 1416
Db     464 AspAspCysSerAlaIleGluCysTyrMetGluGluTyrGlyValThrAlaGlnGluAla 483
Qy    1417 ------------CGGAAGCACGTGAAATGGCTGATAGCGGAGGTGTGGAAGAAGATGAAT 1464
Db     484 TyrAspValPheAsnLysHisValGlu------------SerAlaTrpLysAspValAsn 499
Qy    1465 GCGGAGAGGGTGTCGAAGGATTCTCCATTCGGCAAAGATTTTATAGGATGTGCAGTTGAT 1524
Db     500 LysGluPheLeu---LysProThrGluMetProThrGluValLeuAsnArgSerLeuAsn 518
Qy    1525 TTAGGAAGGATGGCGCAGTTGATGTACCATAATGGAGATGGGCACGGCACACAACACCCT 1584
Db     519 LeuAlaArgValMetAspValLeuTyrArgGluGlyAspGlyTyrThrTyrValGlyLys 538
Qy    1585 ATTATACATCAACAAATGACCAGAACCTTATTCGAGCCCTTTGCA 1629
Db     539 AlaAlaLysGlyGlyIleThrSerLeuLeuIleGluProValAla 553
```

RESULT 9
S68365
(+)-delta-cadinene synthase isozyme XC1 - Gossypium arboreum
C;Species: Gossypium arboreum
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 29-Sep-1999
C;Accession: S68365
R;Chen, X.Y.; Chen, Y.; Heinstein, P.; Davisson, V.J.
Arch. Biochem. Biophys. 324, 255-266, 1995
A;Title: Cloning, expression, and characterization of (+)-delta-cadinene synthase: a ca
A;Reference number: S68365; MUID:96132653; PMID:8554317
A;Accession: S68365
A;Molecule type: mRNA
A;Residues: 1-554 <CHE>
A;Cross-references: EMBL:U23206; NID:g1045311; PIDN:AAA93064.1; PID:g1045312
A;Experimental source: cultivar Nanking
C;Superfamily: vetispiradiene synthase 1
C;Keywords: phytoalexin biosynthesis

```
Alignment Scores:
Pred. No.:                3.15e-49     Length:          554
Score:                    776.00       Matches:         178
Percent Similarity:       52.07%       Conservative:    111
Best Local Similarity:    32.07%       Mismatches:      232
Query Match:              26.69%       Indels:          34
DB:                       2            Gaps:            7

US-09-938-956-6 (1-1632) x S68365 (1-554)

Qy       1 ATGAGACGATCCGGAAACTACAACCCTTCTCGTTGG------------------------ 36
Db      21 MetArgProLysAlaAspPheGlnProSerIleTrpGlyAspLeuPheLeuAsnCysPro 40
Qy      37 GATGTCAACTTCATCCAATCGCTTCTCAGTGACTATAAGGAGGACAAACACGTGATTAGG 96
Db      41 AspLysAsnIleAspAlaGluThrGluLysArgHisGlnGlnLeuLysGluGluValArg 60
Qy      97 GCTTCTGAGCTGGTCACTTTGGTGAAGATGGAACTGGAGAAAGAAACGGATCAAATTCGA 156
Db      61 ------------------------LysMetIleValAlaProMetAlaAsnSerThrGln 72
Qy     157 CAACTTGAGTTGATCGATGACTTGCAGAGGATGGGGCTGTCCGATCATTTCCAAAATGAG 216
Db      73 LysLeuAlaPheIleAspSerValGlnArgLeuGlyValSerTyrHisPheThrLysGlu 92
Qy     217 TTCAAAGAAATCTTGTCCTCTATATATCTCGACCATCACTATTACAAGAACCCTTTTCCA 276
Db      93 IleGluAspGluLeuGluAsnIleTyrHisAsnAsnAsn--------------------- 105
Qy     277 AAAGAAGAAAGGGATCTCTACTCCACATCTCTTGCATTTAGGCTCCTCAGAGAACATGGT 336
Db     106 AspAlaGluAsnAspLeuTyrThrThrSerIleArgPheArgLeuLeuArgGluHisGly 125
Qy     337 TTTCAAGTCGCACAAGAGGTATTCGATAGTTTCAAGAACGAGGAGGGTGAGTTCAAAGAA 396
Db     126 TyrAsnValSerCysAspValPheAsnLysPheLysAspGluGlnGlyAsnPheLysSer 145
Qy     397 AGCCTTAGCGACGACACCAGAGGATTGTTGCAACTGTATGAAGCTTCCTTTCTGTTGACG 456
Db     146 SerValThrSerAspValArgGlyLeuLeuGluLeuTyrGlnAlaSerTyrLeuArgVal 165
Qy     457 GAAGGCGAAACCACGCTCGAGTCAGCGAGGGAATTCGCCACCAAATTTTTGGAGGAAAAA 516
Db     166 HisGlyGluAspIleLeuAspGluAlaIleSerPheThrThrHisHisLeuSerLeuAla 185
Qy     517 GTGAACGAGGGTGGTGTTGATGGCGACCTTTTAACAAGAATCGCATATTCTTTGGACATC 576
Db     186 Val------AlaSerLeuAspHisProLeuSerGluGluValSerHisAlaLeuLysGln 203
Qy     577 CCTCTTCATTGGAGGATTAAAAGGCCAAATGCACCTGTGTGGATCGAATGGTATAGGAAG 636
Db     204 SerIleArgArgGlyLeuProArgValGluAlaArgHisTyrLeuSerValTyrGlnAsp 223
Qy     637 AGGCCCGACATGAATCCAGTAGTGTTGGAGCTTGCCATACTCGACTTAAATATTGTTCAA 696
Db     224 IleGluSerHisAsnLysAlaLeuLeuGluPheAlaLysIleAspPheAsnMetLeuGln 243
Qy     697 GCACAATTTCAAGAAGAGCTCAAAGAATCCTTCAGGTGGTGGAGAAATACTGGGTTTGTT 756
Db     244 PheLeuHisArgLysGluLeuSerGluIleCysArgTrpTrpLysAspLeuAspPheGln 263
```

```
Qy    757 GAGAAGCTGCCCTTCGCAAGGGATAGACTGGTGGAATGCTACTTTTGGAATACTGGGATC 816
Db    264 ArgLysLeuProTyrAlaArgAspArgValValGluGlyTyrPheTrpIleSerGlyVal 283

Qy    817 ATCGAGCCACGTCAGCATGCAAGTGCAAGGATAATGATGGGCAAAGTCAACGCTCTGATT 876
Db    284 TyrPheGluProGlnTyrSerLeuGlyArgLysMetLeuThrLysValIleAlaMetAla 303

Qy    877 ACGGTGATCGATGATATTTATGATGTCTATGGCACCTTAGAAGAACTCGAACAATTCACT 936
Db    304 SerIleValAspAspThrTyrAspSerTyrAlaThrTyrGluGluLeuIleProTyrThr 323

Qy    937 GACCTCATTCGAAGATGGGATATAAACTCAATCGACCAACTTCCCGATTACATGCAACTG 996
Db    324 AsnAlaIleGluArgTrpAspIleLysCysIleAspGluIleProGluTyrMetLysPro 343

Qy    997 TGCTTTCTTGCACTCAACAACTTCGTCGATGATACATCGTACGATGTTATGAAGGAGAAA 1056
Db    344 SerTyrLysAlaLeuLeuAspValTyrGluGluMetValGlnLeuValAlaGluHisGly 363

Qy   1057 GGCGTCAACGTTATACCCTACCTGCGGCAATCGTGGGTTGATTTGGCGGATAAGTATATG 1116
Db    364 ArgGlnTyrArgValGluTyrAlaLysAsnAlaMetIleArgLeuAlaGlnSerTyrLeu 383

Qy   1117 GTAGAGGCACGGTGGTTCTACGGCGGGCACAAACCAAGTTTGGAAGAGTATTTGGAGAAC 1176
Db    384 ValGluAlaLysTrpThrLeuGlnAsnTyrLysProSerPheGluGluPheLysAlaAsn 403

Qy   1177 TCATGGCAGTCGATAAGTGGGCCCTGTATGTTAACGCACATATTCTTCCGAGTAACAGAT 1236
Db    404 AlaLeuProThrCysGlyTyrAlaMetLeuAlaIleThrSerPheValGlyMetGlyAsp 423

Qy   1237 TCGTTCACAAAGGAGACCGTCGACAGTTTGTACAAATACCACGATTTAGTTCGTTGGTCA 1296
Db    424 IleValThrProGluThrPheLysTrpAlaAlaSerAspProLysIleIleGlnAlaSer 443

Qy   1297 TCCTTCGTTCTGCGGCTTGCTGATGATTTGGGAACCTCGGTGGAAGAGGTGAGCAGAGGG 1356
Db    444 ThrIleIleCysArgPheMetAspAspValAlaGluHisLysPheLysHisArgArgGlu 463

Qy   1357 GATGTGCCGAAATCACTTCAGTGCTACATGAGTGACTACAATGCATCGGAGGCGGAGGCG 1416
Db    464 AspAspCysSerAlaIleGluCysTyrMetGluGluTyrGlyValThrAlaGlnGluAla 483

Qy   1417 ------------CGGAAGCACGTGAAATGGCTGATAGCGGAGGTGTGGAAGAAGATGAAT 1464
Db    484 TyrAspValPheAsnLysHisValGlu------------SerAlaTrpLysAspLeuAsn 499

Qy   1465 GCGGAGAGGGTGTCGAAGGATTCTCCATTCGGCAAAGATTTTATAGGATGTGCAGTTGAT 1524
Db    500 GlnGluPheLeu---LysProThrGluMetProThrGluValLeuAsnArgSerLeuAsn 518

Qy   1525 TTAGGAAGGATGGCGCAGTTGATGTACCATAATGGAGATGGGCACGGCACACAACACCCT 1584
Db    519 LeuAlaArgValMetAspValLeuTyrArgGluGlyAspGlyTyrThrTyrValGlyLys 538

Qy   1585 ATTATACATCAACAAATGACCAGAACCTTATTCGAGCCCTTTGCA 1629
Db    539 AlaAlaLysGlyGlyIleThrSerLeuLeuIleGluProIleAla 553
```

RESULT 10
T06266
germacrene C synthase - tomato
C;Species: Lycopersicon esculentum (tomato)
C;Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 11-May-2000
C;Accession: T06266
R;Colby, S.M.; Crock, J.; Dowdle-Rizzo, B.; Lemaux, P.G.; Croteau, R.
Proc. Natl. Acad. Sci. U.S.A. 95, 2216-2221, 1998
A;Title: Germacrene C synthase from Lycopersicon esculentum cv. VFNT Cherry tomato: cDNA
A;Reference number: Z15576; MUID:98151492; PMID:9482865
A;Accession: T06266
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-548 <COL>
A;Cross-references: EMBL:AF035631; NID:g2967688; PIDN:AAC39432.1; PID:g2967689
A;Experimental source: cultivar VFNT
C;Superfamily: vetispiradiene synthase 1

```
Alignment Scores:
Pred. No.:              3.92e-46      Length:         548
Score:                  734.00        Matches:        177
Percent Similarity:     50.72%        Conservative:   106
Best Local Similarity:  31.72%        Mismatches:     233
Query Match:            25.24%        Indels:         42
DB:                     2             Gaps:           11
```

US-09-938-956-6 (1-1632) x T06266 (1-548)

```
Qy      4 AGACGATCCGGAAACTACAACCCTTCTCGTTGGGATGTCAACTTCATCCAATCGCTT--- 60
Db     10 ArgProLeuAlaAsnPheHisProSerValTrpGlyTyrHisPheLeuSerTyrThrHis 29

Qy     61 ---------------------------CTCAGTGACTATAAGGAGGACAAACACGTGATT 93
Db     30 GluIleThrAsnGlnGluLysValGluValAspGluTyrLysGlu--------------- 44

Qy     94 AGGGCTTCTGAGCTGGTCACTTTGGTGAAGATGGAACTGGAGAAAGAAACGGATCAAATT 153
Db     45 ------------------ThrIleArgLysMetLeuValGluThrCysAspAsnSerThr 58

Qy    154 CGACAACTTGAGTTGATCGATGACTTGCAGAGGATGGGGCTGTCCGATCATTTCCAAAAT 213
Db     59 GlnLysLeuValLeuIleAspAlaMetGlnArgLeuGlyValAlaTyrHisPheAspAsn 78

Qy    214 GAGTTCAAAGAAATCTTGTCCTCTATATATCTCGACCATCACTATTACAAGAACCCTTTT 273
Db     79 GluIleGluThrSerIleGlnAsnIlePhe---AspAlaSerSerLysGlnAsn------ 95

Qy    274 CCAAAAGAAGAAAGGGATCTCTACTCCACATCTCTTGCATTTAGGCTCCTCAGAGAACAT 333
Db     96 ---AspAsnAspAsnAsnLeuTyrValValSerLeuArgPheArgLeuValArgGlnGln 114

Qy    334 GGTTTTCAAGTCGCACAAGAGGTATTCGATAGTTTCAAGAACGAGGAGGGTGAGTTCAAA 393
Db    115 GlyHisTyrMetSerSerAspValPheLysGlnPheThrAsnGlnAspGlyLysPheLys 134

Qy    394 GAAAGCCTTAGCGACGACACCAGAGGATTGTTGCAACTGTATGAAGCTTCCTTTCTGTTG 453
Db    135 GluThrLeuThrAsnAspValGlnGlyLeuLeuSerLeuTyrGluAlaSerHisLeuArg 154

Qy    454 ACGGAAGGCGAAACCACGCTCGAGTCAGCGAGGGAATTCGCCACCAAATTTTTGGAGGAA 513
Db    155 ValArgAsnGluGluIleLeuGluGluAlaLeuThrPheThrThrThrHisLeuGluSer 174

Qy    514 AAAGTG---AACGAGGGTGGTGTTGATGGCGACCTTTTAACAAGAATCGCATATTCTTTG 570
Db    175 IleValSerAsnLeuSerAsnAsnAsnAsnSerLeuLysValGluValGlyGluAlaLeu 194

Qy    571 GACATCCCTCTTCATTGGAGGATTAAAAGGCCAAATGCACCTGTGTGGATCGAATGGTAT 630
Db    195 ThrGlnProIleArgMetThrLeuProArgMetGlyAlaArgLysTyrIleSerIleTyr 214

Qy    631 AGGAAGAGGCCCGACATGAATCCAGTAGTGTTGGAGCTTGCCATACTCGACTTAAATATT 690
Db    215 GluAsnAsnAspAlaHisHisHisLeuLeuLeuLysPheAlaLysLeuAspPheAsnMet 234

Qy    691 GTTCAAGCACAATTTCAAGAAGAGCTCAAAGAATCCTTCAGGTGGTGGAGAAATACTGGG 750
Db    235 LeuGlnLysPheHisGlnArgGluLeuSerAspLeuThrArgTrpTrpLysAspLeuAsp 254

Qy    751 TTTGTTGAGAAGCTGCCCTTCGCAAGGGATAGACTGGTGGAATGCTACTTTTGGAATACT 810
Db    255 PheAlaAsnLysTyrProTyrAlaArgAspArgLeuValGluCysTyrPheTrpIleLeu 274

Qy    811 GGGATCATCGAGCCACGTCAGCATGCAAGTGCAAGGATAATGATGGGCAAAGTCAACGCT 870
```

```
Db      275 GlyValTyrPheGluProLysTyrSerArgAlaArgLysMetMetThrLysValLeuAsn 294

Qy      871 CTGATTACGGTGATCGATGATATTTATGATGTCTATGGCACCTTAGAAGAACTCGAACAA 930

Db      295 LeuThrSerIleIleAspAspThrPheAspAlaTyrAlaThrPheAspGluLeuValThr 314

Qy      931 TTCACTGACCTCATTCGAAGATGGGATATAAACTCAATCGACCAACTTCCCGATTACATG 990

Db      315 PheAsnAspAlaIleGlnArgTrpAspAlaAsnAlaIleAspSerIleGlnProTyrMet 334

Qy      991 CAACTGTGCTTTCTTGCACTCAACAACTTCGTCGATGATACATCGTACGATGTTATGAAG 1050

Db      335 ArgProAlaTyrGlnAlaLeuLeuAspIleTyrSerGluMetGluGlnValLeuSerLys 354

Qy     1051 GAGAAAGGCGTCAACGTTATACCCTACCTGCGGCAATCGTGGGTTGATTTGGCGGATAAG 1110

Db      355 GluGlyLysLeuAspArgValTyrTyrAlaLysAsnGluMetLysLysLeuValArgAla 374

Qy     1111 TATATGGTAGAGGCACGGTGGTTCTACGGC---GGGCACAAACCAAGTTTGGAAGAGTAT 1167

Db      375 TyrPheLysGluThrGlnTrpLeuAsnAspCysAspHisIleProLysTyrGluGluGln 394

Qy     1168 TTGGAGAACTCA---------------TGGCAGTCGATAAGTGGGCCCTGTATGTTAACG 1212

Db      395 ValGluAsnAlaIleValSerAlaGlyTyrMetMetIleSerThrThrCysLeuVal--- 413

Qy     1213 CACATATTCTTCCGAGTAACAGATTCGTTCACAAAGGAGACCGTCGACAGTTTGTACAAA 1272

Db      414 ------------GlyIleGluGluPheIleSerHisGluThrPheGluTrpLeuMetAsn 429

Qy     1273 TACCACGATTTAGTTCGTTGGTCATCCTTCGTTCTGCGGCTTGCTGATGATTTGGGAACC 1332

Db      430 GluSerValIleValArgAlaSerAlaLeuIleAlaArgAlaMetAsnAspIleValGly 449

Qy     1333 TCGGTGGAAGAGGTGAGCAGAGGGGATGTGCCGAAATCACTTCAGTGCTACATGAGTGAC 1392

Db      450 HisGluAspGluGlnGluArgGlyHisValAlaSerLeuIleGluCysTyrMetLysAsp 469

Qy     1393 TACAATGCATCGGAGGCGGAGGCGCGGAAGCACGTGAAATGGCTG------ATAGCGGAG 1446

Db      470 TyrGlyAlaSerLysGlnGluThr------TyrIleLysPheLeuLysGluValThrAsn 487

Qy     1447 GTGTGGAAGAAGATGAATGCGGAGAGGGTGTCGAAGGATTCTCCATTCGGCAAAGATTTT 1506

Db      488 AlaTrpLysAspIleAsn---LysGlnPheSerArgProThrGluValProMetPheVal 506

Qy     1507 ATAGGATGTGCAGTTGATTTAGGAAGGATGGCGCAGTTGATGTACCATAATGGAGATGGG 1566

Db      507 LeuGluArgValLeuAsnLeuThrArgValAlaAspThrLeuTyrLysGluLysAspThr 526

Qy     1567 CACGGCACACAACACCCTATTATACATCAACAAATGACCAGAACCTTATTCGAG 1620

Db      527 TyrSerThrAlaLysGlyLysLeuLysAsnMetIleAsnProIleLeuIleGlu 544
```

RESULT 11
T06265
germacrene C synthase, epidermal - tomato
C;Species: Lycopersicon esculentum (tomato)
C;Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 11-May-2000
C;Accession: T06265
R;Colby, S.M.; Crock, J.; Dowdle-Rizzo, B.; Lemaux, P.G.; Croteau, R.
Proc. Natl. Acad. Sci. U.S.A. 95, 2216-2221, 1998
A;Title: Germacrene C synthase from Lycopersicon esculentum cv. VFNT Cherry tomato: cDNA
A;Reference number: Z15576; MUID:98151492; PMID:9482865
A;Accession: T06265
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-548 <COL>
A;Cross-references: EMBL:AF035630; NID:g2967686; PIDN:AAC39431.1; PID:g2967687
A;Experimental source: cultivar VFNT
C;Superfamily: vetispiradiene synthase 1

```
Alignment Scores:
Pred. No.:                2.14e-45     Length:          548
Score:                    724.00       Matches:         175
Percent Similarity:       51.39%       Conservative:    103
Best Local Similarity:    32.35%       Mismatches:      217
Query Match:              24.90%       Indels:          46
DB:                       2            Gaps:            12

US-09-938-956-6 (1-1632) x T06265 (1-548)

Qy        4 AGACGATCCGGAAACTACAACCCTTCTCGTTGGGATGTCAACTTCATCCAATCGCTT--- 60

Db       10 ArgProLeuAlaAsnPheHisProSerValTrpGlyTyrHisPheLeuSerTyrThrHis 29

Qy       61 --------------------------CTCAGTGACTATAAGGAGGACAAACACGTGATT 93

Db       30 GluIleThrAsnGlnGluLysValGluValAspGluTyrLysGlu--------------- 44

Qy       94 AGGGCTTCTGAGCTGGTCACTTTGGTGAAGATGGAACTGGAGAAAGAAACGGATCAAATT 153

Db       45 -----------------ThrIleArgLysMetLeuValGluThrCysAspAsnSerThr 58

Qy      154 CGACAACTTGAGTTGATCGATGACTTGCAGAGGATGGGGCTGTCCGATCATTTCCAAAAT 213

Db       59 GlnLysLeuValLeuIleAspAlaMetGlnArgLeuGlyValAlaTyrHisPheAspAsn 78

Qy      214 GAGTTCAAAGAAATCTTGTCCTCTATATATCTCGACCATCACTATTACAAGAACCCTTTT 273

Db       79 GluIleGluThrSerIleGlnAsnIlePhe---AspAlaSerSerLysGlnAsn------ 95

Qy      274 CCAAAAGAAGAAAGGGATCTCTACTCCACATCTCTTGCATTTAGGCTCCTCAGAGAACAT 333

Db       96 ---AspAsnAspAsnAsnLeuTyrValValSerLeuArgPheArgLeuValArgGlnGln 114

Qy      334 GGTTTTCAAGTCGCACAAGAGGTATTCGATAGTTTCAAGAACGAGGAGGGTGAGTTCAAA 393

Db      115 GlyHisTyrMetSerSerAspValPheLysGlnPheThrAsnGlnAspGlyLysPheLys 134

Qy      394 GAAAGCCTTAGCGACGACACCAGAGGATTGTTGCAACTGTATGAAGCTTCCTTTCTGTTG 453

Db      135 GluThrLeuThrAsnAspValGlnGlyLeuLeuSerLeuTyrGluAlaSerHisLeuArg 154

Qy      454 ACGGAAGGCGAAACCACGCTCGAGTCAGCGAGGGAATTCGCCACCAAATTTTTGGAGGAA 513

Db      155 ValArgAsnGluGluIleLeuGluGluAlaLeuThrPheThrThrThrHisLeuGluSer 174

Qy      514 AAAGTG---AACGAGGGTGGTGTTGATGGCGACCTTTTAACAAGAATCGCATATTCTTTG 570

Db      175 IleValSerAsnLeuSerAsnAsnAsnAsnSerLeuLysValGluValGlyGluAlaLeu 194

Qy      571 GACATCCCTCTTCATTGGAGGATTAAAAGGCCAAATGCACCTGTGTGGATCGAATGGTAT 630

Db      195 ThrGlnProIleArgMetThrLeuProArgMetGlyAlaArgLysTyrIleSerIleTyr 214

Qy      631 AGGAAGAGGCCCGACATGAATCCAGTAGTGTTGGAGCTTGCCATACTCGACTTAAATATT 690

Db      215 GluAsnAsnAspAlaHisHisHisLeuLeuLeuLysPheAlaLysLeuAspPheAsnMet 234

Qy      691 GTTCAAGCACAATTTCAAGAAGAGCTCAAAGAATCCTTCAGGTGGTGGAGAAATACTGGG 750

Db      235 LeuGlnLysPheHisGlnArgGluLeuSerAspLeuThrArgTrpTrpLysAspLeuAsp 254

Qy      751 TTTGTTGAGAAGCTGCCCTTCGCAAGGGATAGACTGGTGGAATGCTACTTTTGGAATACT 810

Db      255 PheAlaAsnLysTyrProTyrAlaArgAspArgLeuValGluCysTyrPheTrpIleLeu 274

Qy      811 GGGATCATCGAGCCACGTCAGCATGCAAGTGCAAGGATAATGATGGGCAAAGTCAACGCT 870

Db      275 GlyValTyrPheGluProLysTyrSerArgAlaArgLysMetMetThrLysValLeuAsn 294

Qy      871 CTGATTACGGTGATCGATGATATTTATGATGTCTATGGCACCTTAGAAGAACTCGAACAA 930

Db      295 LeuThrSerIleIleAspAspThrPheAspAlaTyrAlaThrPheAspGluLeuValThr 314
```

```
Qy     931 TTCACTGACCTCATTCGAAGATGGGATATAAACTCAATCGACCAACTTCCCGATTACATG 990
Db     315 PheAsnAspAlaIleGlnArgTrpAspAlaAsnAlaIleAspSerIleGlnProTyrMet 334

Qy     991 CAACTGTGCTTTCTTGCACTCAACAACTTCGTCGATGATACATCGTACGATGTTATGAAG 1050
Db     335 ArgProAlaTyrGlnAlaLeuLeuAspIleTyrSerGluMetGluGlnValLeuSerLys 354

Qy    1051 GAGAAAGGCGTCAACGTTATACCCTACCTGCGGCAATCGTGGGTTGATTTGGCGGATAAG 1110
Db     355 GluGlyLysLeuAspArgValTyrTyrAlaLysAsnGluMetLysLysLeuValArgAla 374

Qy    1111 TATATGGTAGAGGCACGGTGGTTCTACGGC---GGCACAAACCAAGTTTGGAAGAGTAT 1167
Db     375 TyrPheLysGluThrGlnTrpLeuAsnAspCysAspHisIleProLysTyrGluGluGln 394

Qy    1168 TTGGAGAACTCA---------------TGGCAGTCGATAAGTGGGCCCTGTATGTTAACG 1212
Db     395 ValGluAsnAlaIleValSerAlaGlyTyrMetMetIleSerThrThrCysLeuVal--- 413

Qy    1213 CACATATTCTTCCGAGTAACAGATTCGTTCACAAAGGAGACCGTCGACAGTTTGTACAAA 1272
Db     414 ------------GlyIleGluGluPheIleSerHisGluThrPheGluTrpLeuMetAsn 429

Qy    1273 TACCACGATTTAGTTCGTTGGTCATCCTTCGTTCTGCGGCTTGCTGATGATTTGGGAACC 1332
Db     430 GluSerValIleValArgAlaSerAlaLeuIleAlaArgAlaMetAsnAspIleValGly 449

Qy    1333 TCGGTGGAAGAGGTGAGCAGAGGGGATGTGCCGAAATCACTTCAGTGCTACATGAGTGAC 1392
Db     450 HisGluAspGluGlnGluArgGlyHisValAlaSerLeuIleGluCysTyrMetLysAsp 469

Qy    1393 TACAATGCATCGGAGGCGGAGGCGCGGAAGCACGTGAAATGGCTG------ATAGCGGAG 1446
Db     470 TyrGlyAlaSerLysGlnGluThr------TyrIleLysPheLeuLysGluValThrAsn 487

Qy    1447 GTGTGGAAGAAGATGAATGCGGAG------AGGGTGTCGAAGGATTCTCCATTCGGCAAA 1500
Db     488 AlaTrpLysAspIleAsnLysGlnPhePheArgProThrGluValProMetPheValLeu 507

Qy    1501 GATTTTATAGGATGTGCAGTTGATTTAGGAAGGATGGCGCAGTTGATGTACCATAATGGA 1560
Db     508 GluArgVal---------LeuAsnLeuThrArgValAlaAspThrLeuTyrLysGluLys 524

Qy    1561 GAT 1563
Db     525 Asp 525
```

RESULT 12
E96723
hypothetical protein F20P5.19 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001
C;Accession: E96723
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C. C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: E96723
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-632 <STO>
A;Cross-references: GB:AE005173; NID:g2194130; PIDN:AAB61105.1; GSPDB:GN00141
C;Genetics:
A;Gene: F20P5.19
A;Map position: 1
C;Superfamily: vetispiradiene synthase 1

```
Alignment Scores:
Pred. No.:              6.2e-38     Length:          632
Score:                  623.00      Matches:         169
Percent Similarity:     47.05%      Conservative:    94
Best Local Similarity:  30.23%      Mismatches:      232
Query Match:            21.42%      Indels:          64
DB:                     2           Gaps:            11

US-09-938-956-6 (1-1632) x E96723 (1-632)

Qy      25 CCTTCTCGTTGGGATGTCAACTTCATCCAATCGCTTCTCAGTGACTATAAGGAGGACAAA 84
Db      73 ProSerLeuTrpGlyHisHisPheLeuSerAlaSerValAsnGlnThrGluMetAspAsp 92

Qy      85 CACGTGATTAGGGCTTCTGAGCTGGTCACTTTGGTGAAGATGGAACTGGAG------AAA 138
Db      93 LeuTrpArgGlnIleGluAlaLeuLysProIleValAsnAlaMetLeuLeuProCysAsn 112

Qy     139 GAAACGGATCAAATTCGACAACTTGAGTTGATCGATGACTTGCAGAGGATGGGGCTGTCC 198
Db     113 GlyAlaAspAlaLysLysIleThrCysPheIleHisThrLeuValSerLeuGlyValSer 132

Qy     199 GATCATTTCCAAAATGAGTTCAAAGAAATCTTGTCCTCTATATATCTCGACCATCACTAT 258
Db     133 TyrHisPheGluGluLysIleValGluPheLeu--------------------------- 143

Qy     259 TACAAGAACCCTTTTCCAAAA------------------------GAAGAAAGGGATCTC 294
Db     144 ---LysAspAlaPheGluAsnIleGluAspMetIleIleAspCysLysGluAspAspLeu 162

Qy     295 TACTCCACATCTCTTGCATTTAGGCTCCTCAGAGAACATGGTTTTCAAGTCGCACAAGAG 354
Db     163 TyrThrValSerIleIlePheArgValPheArgLeuTyrGlyHisTyrIleThrProGlu 182

Qy     354 ------------------------------------------------------------ 354
Db     183 LeuHisIleThrIleSerTyrTyrPhePhePheLeuIlePheHisThrCysMetCysVal 202

Qy     355 ---GTATTCGATAGTTTCAAGAACGAGGAGGGTGAGTTCAAAGAAAGCCTTAGCGACGAC 411
Db     203 AspIlePheAsnArgPheLysGlyAspAspGlyAsnPheLysLysCysLeuAsnAspAsp 222

Qy     412 ACCAGAGGATTGTTGCAACTGTATGAAGCTTCCTTTCTGTTGACGGAAGGCGAAACCACG 471
Db     223 ValArgGlyMetLeuSerPheTyrGluAlaSerHisPheGlyThrThrThrGluAspIle 242

Qy     472 CTCGAGTCAGCGAGGGAATTCGCCACCAAATTTTTGGAG---GAAAAAGTGAACGAGGGT 528
Db     243 LeuGluGluAlaMetSerPheThrGlnLysHisLeuGluLeuPheLeuValGlyGluLys 262

Qy     529 GGTGTTGATGGCGACCTTTTAACAAGA---ATCGCATATTCTTTGGACATCCCTCTTCAT 585
Db     263 AlaLysHisTyrProHisIleThrLysLeuIleGlnAlaAlaLeuTyrIleProGlnAsn 282

Qy     586 TGGAGGATTAAAAGGCCAAATGCACCTGTGTGGATCGAATGGTATAGGAAGAGGCCCGAC 645
Db     283 PheAsnLeuGluIleLeuValAlaArgGluTyrIleAspPheTyrGluLeuGluThrAsp 302

Qy     646 ATGAATCCAGTAGTGTTGGAGCTTGCCATACTCGACTTAAATATTGTTCAAGCACAATTT 705
Db     303 HisAsnGluMetLeuLeuLysLeuAlaLysLeuAsnPheArgPheLeuGlnLeuGlnTyr 322

Qy     706 CAAGAAGAGCTCAAAGAATCCTTCAGGTGGTGGAGAAATACTGGGTTTGTTGAGAAGCTG 765
Db     323 IleGlnAspLeuLysThrLeuThrThrTrpTrpLysGluLeuAspLeuValSerLysIle 342

Qy     766 CCC---TTCGCAAGGGATAGACTGGTGGAATGCTACTTTTGGAATACTGGGATCATCGAG 822
Db     343 ProValTyrPheArgGluArgLeuAlaGluProTyrPheTrpAlaThrGlyIleTyrTyr 362
```

```
Qy     823 CCACGTCAGCATGCAAGTGCAAGGATAATGATGGGCAAAGTCAACGCTCTGATTACGGTG 882
Db     363 GluProGlnTyrSerAlaAlaArgIleMetLeuAlaLysSerIleIleLeuValAspIle 382

Qy     883 ATCGATGATATTTATGATGTCTATGGCACCTTAGAAGAACTCGAACAATTCACTGACCTC 942
Db     383 ValAspAsnThrPheAspValTyrGlyThrIleAspGluValLysSerLeuValGlnAla 402

Qy     943 ATTCGAAGATGGGATATAAACTCAATCGACCAACTTCCCGATTACATGCAACTGTGCTTT 1002
Db     403 IleGluArgTrpAspSerAspAlaValAspValLeuProAspTyrLeuLysValValPhe 422

Qy    1003 CTTGCACTCAACAACTTCGTCGATGATACATCGTACGATGTTATGAAGGAGAAAGGCGTC 1062
Db     423 -----------------------ArgThrThrPheAspLeuPheLysGluLeuGluGlu 434

Qy    1063 AACGTT-----------------------ATACCCTACCTGCGGCAATCGTGGGTTGAT 1098
Db     435 TyrValSerSerGluAlaArgSerPheThrMetGlnTyrAlaTyrGluGlnLeuArgIle 454

Qy    1099 TTGGCGGATAAGTATATGGTAGAGGCACGGTGGTTCTACGGCGGGCACAAACCAAGTTTG 1158
Db     455 LeuMetLysGlyTyrLeuGlnGluAlaGluTrpSerAsnArgGlyHisLeuProSerHis 474

Qy    1159 GAAGAGTATTTGGAGAACTCATGGCAGTCGATAAGTGGGCCCTGTATGTTAACGCACATA 1218
Db     475 GluGluTyrIleGluValGlyValAlaSerThrAlaGlyGluValLeuLeuAlaMetThr 494

Qy    1219 TTCTTCCGAGTAACAGATTCGTTCACAAAGGAGACCGTCGACAGTTTGTACAAATACCAC 1278
Db     495 PheIleProMetGlyAspAlaAlaGlyValGlyValTyrGluTrpLeuArgSerArgPro 514

Qy    1279 GATTTAGTTCGTTGGTCATCCTTCGTTCTGCGGCTTGCTGATGATTTGGGAACCTCGGTG 1338
Db     515 LysLeuThrHisAlaLeuPheValLysSerArgLeuArgAspAspIleAlaThrTyrLys 534

Qy    1339 GAAGAGGTGAGCAGAGGGGATGTGCCGAAATCACTTCAGTGCTACATGAGTGACTACAAT 1398
Db     535 GluGluMetLysArgGlyAspValCysAsnGlyIleAsnCysTyrThrLysGlnHisLys 554

Qy    1399 GCATCGGAGGCGGAGGCGCGGAAGCACGTGAAATGGCTGATAGCGGAGGTGTGGAAGAAG 1458
Db     555 ValSerGluGluGluAlaCysIleGluPheGluLysLysThrAsnHisMetSerLysVal 574

Qy    1459 ATGAATGCGGAG------AGGGTGTCGAAGGATTCTCCATTCGGCAAAGATTTTATAGGA 1512
Db     575 MetAsnGluGluPheLeuLysAlaAlaLysPheIleProLeu------HisIleLeuArg 592

Qy    1513 TGTGCAGTTGATTTAGGAAGGATGGCGCAGTTGATGTACCATAATGGAGATGGGCAC 1569
Db     593 ProValLeuAsnTyrGlyArgLeuAlaAspValCysTyrLysTyrGlyAspGlyTyr 611
```

RESULT 13
G86443
probable terpene synthase type protein [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C;Accession: G86443
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C. C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: G86443
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-608 <STO>
A;Cross-references: GB:AE005172; NID:g10092429; PIDN:AAG12833.1; GSPDB:GN00141
C;Genetics:
A;Map position: 1
C;Superfamily: vetispiradiene synthase 1

```
Alignment Scores:
Pred. No.:                3.05e-36     Length:         608
Score:                    600.00       Matches:        154
Percent Similarity:       48.11%       Conservative:   113
Best Local Similarity:    27.75%       Mismatches:     250
Query Match:              20.63%       Indels:         38
DB:                       2            Gaps:           8

US-09-938-956-6 (1-1632) x G86443 (1-608)

Qy      16 AACTACAACCCTTCTCGTTGGGATGTCAACTTCATCCAATCGCTTCTCAGT--------- 66
Db      69 HisPheSerProSerLeuTrpGlyAspHisPheLeuSerValSerLeuAsnArgAlaVal 88

Qy      67 GACTATAAGGAGGACAAACACGTGATTAGGGCTTCTGAGCTGGTCACTTTGGTGAAGATG 126
Db      89 GluPheAspGluLeuGluArgGluIleGluThrThrLysProLeuAlaArgAspLysLeu 108

Qy     127 GAACTGGAGAAAGAAACGGATCAAATTCGACAACTTGAGTTGATCGATGACTTGCAGAGG 186
Db     109 ---MetSerSerGluSerSerAspLysGluLysIleArgLeuIleHisLeuLeuValSer 127

Qy     187 ATGGGGCTGTCCGATCATTTCCAAAATGAGTTCAAAGAAATCTTGTCCTCTATATATCTC 246
Db     128 MetGlyIleSerTyrHisPheAspLysGluIleGlnAspIleLeu--------------- 142

Qy     247 GACCATCACTATTACAAGAACCCTTTTCCAAAA-----------------GAAGAAAGG 288
Db     143 ---------------LysHisSerPheThrLysLeuAspAspIleIleValGlyGluAsp 157

Qy     289 GATCTCTACTCCACATCTCTTGCATTTAGGCTCCTCAGAGAACATGGTTTTCAAGTCGCA 348
Db     158 AspLeuGluThrIleSerIleMetPheGluValPheLysLeuTyrGlyHisLysMetSer 177

Qy     349 CAAGAGGTATTCGATAGTTTCAAGAACGAGGAGGGTGAGTTCAAAGAAAGCCTTAGCGAC 408
Db     178 CysAspAlaPheAspArgPheArgGlyAsnAspGlyArgPheLysGluSerLeuValArg 197

Qy     409 GACACCAGAGGATTGTTGCAACTGTATGAAGCTTCCTTTCTGTTGACGGAAGGCGAAACC 468
Db     198 AspPheArgGlyMetLeuGlnLeuPheGluValAlaHisLeuGlyThrProCysGluVal 217

Qy     469 ACGCTCGAGTCAGCGAGGGAATTCGCCACCAAATTTTTGGAGGAAAAAGTG------AAC 522
Db     218 IleMetAspGluAlaLeuSerPheThrArgAsnHisLeuGluSerLeuThrSerGlyAsn 237

Qy     523 GAGGGTGGTGTTGATGGCGACCTTTTAACAAGAATCGCATATTCTTTGGACATCCCTCTT 582
Db     238 AlaSerThrAlaSerProHisLeuLeuLysHisIleGlnAsnSerLeuTyrIleProArg 257

Qy     583 CATTGGAGGATTAAAAGGCCAAATGCACCTGTGTGGATCGAATGGTATAGGAAGAGGCCC 642
Db     258 TyrCysAsnIleGluValLeuValAlaArgGluTyrIleSerTyrTyrGluGlnGluGlu 277

Qy     643 GACATGAATCCAGTAGTGTTGGAGCTTGCCATACTCGACTTAAATATTGTTCAAGCACAA 702
Db     278 GlyHisAspGluIleLeuLeuLysPheAlaLysLeuAsnPheAsnPheCysGlnPheHis 297

Qy     703 TTTCAAGAAGAGCTCAAAGAATCCTTCAGGTGGTGGAGAAATACTGGGTTTGTTGAGAAG 762
Db     298 TyrValGlnGluLeuLysThrLeuThrLysTrpTrpArgAspLeuAspLeuAlaSerLys 317

Qy     763 CTGCCCTTCGCAAGGGATAGACTGGTGGAATGCTACTTTTGGAATACTGGGATCATCGAG 822
Db     318 LeuProTyrIleArgAspArgLeuValGluSerHisLeuValAlaLeuGlyProTyrPhe 337

Qy     823 CCACGTCAGCATGCAAGTGCAAGGATAATGATGGGCAAAGTCAACGCTCTGATTACGGTG 882
```

```
Db        338 GluProHisTyrSerLeuGlyArgIleIleValAlaLysIleAsnMetIleMetValVal 357

Qy        883 ATCGATGATATTTATGATGTCTATGGCACCTTAGAAGAACTCGAACAATTCACTGACCTC 942
              :::||||||    ||||||    |||   ||||||    :::::::::       |||:::
Db        358 ValAspAspThrTyrAspAlaTyrAlaThrLeuProGlnValLysAlaLeuThrGluCys 377

Qy        943 ATTCGAAGATGGGATATAAACTCAATCGACCAACTTCCCGATTACATGCAACTGTGCTTT 1002
              ::::::||||||    |||                 |||:::||||||||||||:::::::::
Db        378 LeuGlnArgTrpSerIleGluValSerAspLysLeuProAspTyrLeuArgIleValLeu 397

Qy       1003 CTTGCACTCAACAACTTCGTCGATGATACATCGTACGATGTTATGAAGGAGAAAGGCGTC 1062
              :::|||                        :::|||||||||   |||
Db        398 GlySerLeu------------------------PheAspValMetGlyGluIleGluArg 409

Qy       1063 AACGTTATACCC------------TACCTGCGGCAATCGTGGGTTGAT------------ 1098
               :::   |||          |||        :::        |||:::
Db        410 GluMetArgProLeuGlyArgLeuTyrArgValLysGlnValValGluLysIleLysIle 429

Qy       1099 TTGGCGGATAAGTATATGGTAGAGGCACGGTGGTTCTACGGCGGGCACAAACCAAGTTTG 1158
              :::             |||         |||:::|||          ||||||      :::
Db        430 IleThrLysAlaTyrGlnGluIleAlaLysTrpAlaArgThrGlyHisValSerThrPhe 449

Qy       1159 GAAGAGTATTTGGAGAACTCATGGCAGTCGATAAGTGGGCCCTGTATGTTAACGCACATA 1218
              :::||||||::::::            :::                          :::
Db        450 AspGluTyrMetLysValGlyValLeuThrAlaGlyMetAlaAspTyrAlaAlaTyrCys 469

Qy       1219 TTCTTCCGAGTAACAGATTCGTTCACAAAGGAGACCGTCGACAGTTTGTACAAATACCAC 1278
              |||       :::    |||          ||||||     :::    |||
Db        470 PheIleGlyMetGluAspIleAsnGluLysGluAlaPheGluTrpLeuAsnSerAsnPro 489

Qy       1279 GATTTAGTTCGTTGGTCATCCTTCGTTCTGCGGCTTGCTGATGATTTGGGAACCTCGGTG 1338
                 :::::::::        :::   :::   |||||||||:::|||:::||||||
Db        490 LeuIleIleLysHisLeuThrAlaMetPheArgLeuAlaAsnAspValGlyThrTyrGlu 509

Qy       1339 GAAGAGGTGAGCAGAGGGGATGTGCCGAAATCACTTCAGTGCTACATGAGTGACTACAAT 1398
                 |||::::::||||||:::|||          |||    |||||||||      |||
Db        510 ThrGluIleAsnArgGlyGluValAlaAsnGlyLeuAsnCysTyrMetLysGlnTyrGly 529

Qy       1399 GCATCGGAGGCGGAGGCGCGGAAGCACGTGAAATGGCTGATAGCGGAGGTGTGGAAGAAG 1458
                 ::::::   ||||||   :::   ::::::   :::                  |||
Db        530 ValThrLysGluGluAlaSerArgGluLeuArgLysMetTyrValTyrArgLysLysVal 549

Qy       1459 ATGAATGCGGAGAGGGTGTCGAAGGATTCTCCATTCGGCAAAGATTTTATAGGATGTGCA 1518
              :::       |||   ::::::                 :::      :::
Db        550 ValValGluGluPheMetAsnSerHisAspHisValProArgGlnValLeuLeuArgCys 569

Qy       1519 GTTGATTTAGGAAGGATGGCGCAGTTGATGTACCATAATGGAGATGGGCACGGCACACAA 1578
              :::::::::   |||:::      :::   |||         |||||||||:::
Db        570 LeuAsnIleAlaArgIlePheAspValPheTyrThrGluGlyAspGlyTyrSerGluPro 589

Qy       1579 CACCCTATTATACATCAACAAATGACCAGAACCTTATTCGAGCCC 1623
                           |||         ||||||         |||
Db        590 LysGlyLysIleGluHisPheMetThrSerLeuTyrLeuHisPro 604
```

RESULT 14
C96642
hypothetical protein T13M11.3 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C;Accession: C96642
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C. C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: C96642
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-421 <STO>
A;Cross-references: GB:AE005173; NID:g4508070; PIDN:AAD21414.1; GSPDB:GN00141
C;Genetics:
A;Gene: T13M11.3
A;Map position: 1

```
Alignment Scores:
Pred. No.:              1.06e-33        Length:         421
Score:                  565.00          Matches:        142
Percent Similarity:     46.28%          Conservative:   82
Best Local Similarity:  29.34%          Mismatches:     166
Query Match:            19.43%          Indels:         94
DB:                     2               Gaps:           10

US-09-938-956-6 (1-1632) x C96642 (1-421)

Qy        166 TTGATCGATGACTTGCAGAGGATGGGGCTGTCCGATCATTTCCAAAATGAGTTCAAAGAA 225
              :::||||||      :::|||   :::|||:::      ||||||:::   |||   ::::::
Db          1 MetIleAspValIleGlnSerLeuGlyIleAspLeuHisPheArgGlnGluIleGluGln 20

Qy        226 ATCTTGTCCTCTATATATCTCGACCATCACTATTACAAGAACCCTTTTCCAAAAGAAGAA 285
                 |||   ||||||   :::         :::
Db         21 ThrLeuHisMetIleTyrLysGluGlyLeuGlnPheAsnGly------------------ 34

Qy        286 AGGGATCTCTACTCCACATCTCTTGCATTTAGGCTCCTCAGAGAACATGGTTTTCAAGTC 345
                 ||||||:::      :::|||   |||||||||||||||:::   |||      |||
Db         35 ---AspLeuHisGluIleAlaLeuArgPheArgLeuLeuArgGlnGluGlyHisTyrVal 53

Qy        346 GCACAAGAGGTATTCGATAGTTTCAAGAACGAGGAGGGTGAGTTCAAAGAAAGCCTTAGC 405
              :::                        |||::::::|||   ||||||:::   :::
Db         54 GlnGlu---------------------AsnLysLysGlyGlyPheLysAspValValLys 66

Qy        406 GACGACACCAGAGGATTGTTGCAACTGTATGAAGCTTCCTTTCTGTTGACGGAAGGCGAA 465
              :::|||   :::||||||   :::|||:::|||||||||   |||      |||||||||
Db         67 AsnAspValLysGlyLeuThrGluLeuPheGluAlaSerGluLeuArgValGluGlyGlu 86

Qy        466 ACCACGCTCGAGTCAGCGAGGGAATTCGCCACCAAATTTTGGAGGAAAAAGTGAACGAG 525
                 ||||||:::   ||||||||||||          |||   |||      :::
Db         87 GluThrLeuAspGlyAlaArgGluPheThrTyrSerArgLeuAsnGluLeuCysSer--- 105

Qy        526 GGTGGTGTTGATGGCGACCTTTTAACAAGAATCGCATATTCTTTGGACATCCCTCTTCAT 585
                 |||   :::         |||      ||||||      |||   |||
Db        106 ---GlyArgGluSerHisGlnLysGlnGluIleMetLysSerLeuAlaGlnProArgHis 124

Qy        586 TGGAGGATT---------------AAAAGGCCAAATGCACCTGTGTGGATC--------- 621
                 :::            ||||||   :::   :::   |||
Db        125 LysThrValArgGlyLeuThrSerLysArgPheThrSerMetIleLysIleAlaGlyGln 144

Qy        622 ---------GAATGGTATAGGAAGAGGCCCGACATGAATCCAGTAGTGTTGGAGCTTGCC 672
                       ||||||   :::                    :::|||   :::|||
Db        145 GluAspProGluTrpLeuGlnSer---------------------LeuLeuArgValAla 157

Qy        673 ATACTCGACTTAAATATTGTTCAAGCACAATTTCAAGAAGAGCTCAAAGAATCCTTCAGG 732
                 :::|||         :::::::::      |||   |||:::   ::::::|||:::
Db        158 GluIleAspSerIleArgLeuLysSerLeuThrGlnGlyGluMetSerGlnThrPheLys 177

Qy        733 TGGTGGAGAAATACTGGGTTTGTTGAGAAGCTGCCCTTCGCAAGGGATAGACTGGTGGAA 792
              ||||||   |||                :::   :::   ||||||   :::   ::::::
Db        178 TrpTrpThrGluLeuGlyLeuGluLysAspValGluLysAlaArgSerGlnProLeuLys 197

Qy        793 TGCTACTTTTGGAATACTGGGATCATCGAGCCACGTCAGCATGCAAGTGCAAGGATAATG 852
                 :::   |||:::      |||::::::                  |||:::
Db        198 TrpHisThrTrpSerMetLysIleLeuGlnAspProThrLeuThrGluGlnArgLeuAsp 217

Qy        853 ATGGGCAAAGTCAACGCTCTGATTACGGTGATCGATGATATTTATGATGTCTATGGCACC 912
              :::   |||      :::|||:::   |||||||||||||||:::||||||||||||
Db        218 LeuThrLysProIleSerLeuValTyrValIleAspAspIlePheAspValTyrGlyGlu 237

Qy        913 TTAGAAGAACTCGAACAATTCACTGACCTCATTCGAAGATGGGATATAAACTCAATCGAC 972
              ||||||||||||         ||||||   ::::::   |||||||||         :::
Db        238 LeuGluGluLeuThrIlePheThrArgValValGluArgTrpAspHisLysGlyLeuLys 257
```

```
Qy        973 CAACTTCCCGATTACATGCAACTGTGCTTTCTTGCACTCAACAACTTCGTCGATGATACA 1032
                 ||||||   |||||||::::::||||||   ||||||:::             :::
Db        258 ThrLeuProLysTyrMetArgValCysPheGluAlaLeuAspMetIleThrThrGluIle 277

Qy       1033 TCGTACGATGTTATGAAGGAGAAAGGCGTCAACGTTATACCCTACCTGCGGCAATCG--- 1089
              |||      :::  |||      |||  |||            |||||||||||
Db        278 SerMetLysIleTyrLysSerHisGlyTrpAsnProThrTyrAlaLeuArgGlnSerVal 297

Qy       1090 ---------------TGGGTTGATTTGGCGGATAAGTATATGGTAGAGGCACGGTGGTTC 1134
                             |||       |||          ::::::|||||||||:::||||||
Db        298 IleIleIleGlnAsnTrpAlaSerLeuCysLysAlaPheLeuValGluAlaLysTrpPhe 317

Qy       1135 TACGGCGGGCACAAACCAAGTTTGGAAGAGTATTTGGAGAACTCATGGCAGTCGATAAGT 1194
                    |||:::   |||:::   ||||||||::::::|||
Db        318 AsnSerGlyTyrLeuProThrThrGluGluTyrMetLysAsn------------------ 331

Qy       1195 GGGCCCTGTATGTTAACGCACATATTCTTCCGAGTAACAGATTCGTTCACAAAGGAGACC 1254

Db        331 ------------------------------------------------------------ 331

Qy       1255 GTCGACAGTTTGTACAAATACCACGATTTAGTTCGTTGGTCATCCTTCGTTCTGCGGCTT 1314

Db        331 ------------------------------------------------------------ 331

Qy       1315 GCTGATGATTTGGGAACCTCGGTGGAAGAGGTGAGCAGAGGGGATGTGCCGAAATCACTT 1374
                                   |||:::|||         |||                :::
Db        332 ------------------GlyValAspGluAsnGlnAspGlyThrAspGlySerTyrVal 345

Qy       1375 CAGTGCTACATGAGTGACTACAATGCATCGGAGGCG---GAGGCGCGGAAGCACGTGAAA 1431
              :::||||||:::::::::|||      |||         |||||||||   ||||||
Db        346 GluCysTyrLeuAsnGluTyrLysGlySerThrValAspGluAlaArgThrHisValAla 365

Qy       1432 TGGCTGATAGCGGAGGTGTGGAAGAAGATGAATGCGGAGAGGGTGTCGAAGGATTCTCCA 1491
                 |||:::      ||||||::::::|||   |||    ::::::            |||
Db        366 GlnLysIleSerArgAlaTrpLysArgLeuAsnArgGluCysLeuAsn---ProCysPro 384

Qy       1492 TTCGGCAAAGATTTTATAGGATGTGCAGTTGATTTAGGAAGGATGGCGCAGTTGATGTAC 1551
              |||   :::   |||               :::::::::   |||      ||||||||
Db        385 PheSerArgSerPheSerLysAlaCysLeuAsnIleAlaArgThrValProLeuMetTyr 404

Qy       1552 CATAATGGAGAT 1563
                       |||
Db        405 SerTyrAspAsp 408
```

RESULT 15
C56118
vetispiradiene synthase 2 - Hyoscyamus muticus (fragment)
C;Species: Hyoscyamus muticus
C;Date: 19-Oct-1995 #sequence_revision 19-Oct-1995 #text_change 29-Sep-1999
C;Accession: C56118
R;Back, K.; Chappell, J.
J. Biol. Chem. 270, 7375-7381, 1995
A;Title: Cloning and bacterial expression of a sesquiterpene cyclase from Hyoscyamus mut
A;Reference number: A56118; MUID:95221394; PMID:7706281
A;Accession: C56118
A;Status: preliminary; nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-350 <BAC>
A;Cross-references: GB:U20190; NID:g763426; PIDN:AAA86339.1; PID:g763427
A;Note: authors translated the codon GAA for residue 181 as Val, and GCA for residue 182
C;Superfamily: vetispiradiene synthase 1

Alignment Scores:

Pred. No.:	1.55e-33	Length:	350
Score:	562.50	Matches:	116
Percent Similarity:	53.64%	Conservative:	68
Best Local Similarity:	33.82%	Mismatches:	140
Query Match:	19.34%	Indels:	19
DB:	2	Gaps:	5

US-09-938-956-6 (1-1632) x C56118 (1-350)

```
Qy        580 CTTCATTGGAGGATTAAAAGGCCAAATGCACCTGTGTGGATCGAATGGTATAGGAAGAGG 639
                ||||||      |||   |||                    :::|||      |||   :::
Db          1 LeuHisLysSerIleProArgValGluThrArgTyrPheIleThrIleTyrGluGluGlu 20

Qy        640 CCCGACATGAATCCAGTAGTGTTGGAGCTTGCCATACTCGACTTAAATATTGTTCAAGCA 699
                       |||   |||:::|||      |||   ||||||   |||::::::|||
Db         21 GluLeuLysAsnAspValLeuLeuArgPheAlaLysLeuAspPheAsnLeuLeuGlnMet 40

Qy        700 CAATTTCAAGAAGAGCTCAAAGAATCCTTCAGGTGGTGGAGAAATACTGGGTTTGTTGAG 759
                 ::::::||||||   |||         ||||||::::::         ||||||
Db         41 LeuHisLysGlnGluLeuThrGluValSerMetTrpTrpLysAspLeuAspPheValThr 60

Qy        760 AAGCTGCCCTTCGCAAGGGATAGACTGGTGGAATGCTACTTTTGGAATACTGGGATCATC 819
                 ||||||:::||||||||||||   ||||||||||||||||||      |||:::
Db         61 ThrLeuProTyrAlaArgAspArgAlaValGluCysTyrPheTrpThrValGlyValTyr 80

Qy        820 GAGCCACGTCAGCATGCAAGTGCAAGGATAATGATGGGCAAAGTCAACGCTCTGATTACG 879
                    |||::::::   ||||||:::|||:::   |||      |||:::|||:::
Db         81 AlaGluProGlnTyrSerGluAlaArgValMetLeuAlaLysThrIleAlaMetIleSer 100

Qy        880 GTGATCGATGATATTTATGATGTCTATGGCACCTTAGAAGAACTCGAACAATTCACTGAC 939
              ::::::||||||   :::|||   ||||||   ::::::|||||||||   :::||||||
Db        101 IleValAspAspThrPheAspAlaTyrGlyIleValLysGluLeuGluValTyrThrAsp 120

Qy        940 CTCATTCGAAGATGGGATATAAACTCAATCGACCAACTTCCCGATTACATGCAACTGTGC 999
                 |||:::||||||||||||   ||||||:::||||||||||||||::::::
Db        121 AlaIleGlnArgTrpAspIleAsnGlnIleAspArgLeuProAspTyrMetLysIleSer 140

Qy       1000 TTTCTTGCACTCAACAACTTCGTCGATGATACATCGTACGATGTTATGAAGGAGAAAGGC 1059
              :::      |||   :::            |||      ::::::   |||:::
Db        141 TyrLysValLeuLeuAspLeuTyrLysAspTyrGluThrGluLeuSerLysAspGlyArg 160

Qy       1060 GTCAACGTTATACCCTACCTGCGGCAATCGTGGGTTGATTTGGCGGATAAGTATATGGTA 1119
                    |||:::   |||      ::::::               |||   |||
Db        161 SerGluValValHisTyrAlaLysGluArgMetLysGluIleValArgAsnTyrPheVal 180

Qy       1120 GAGGCACGGTGGTTCTACGGCGGGCACAAACCAAGTTTGGAAGAGTATTTGGAGAACTCA 1179
              ||||||:::||||||         |||:::   |||   :::   |||||||||   |||
Db        181 GluAlaLysTrpPheIleGluGlyTyrMetProProValSerGluTyrLeuAsnAsnArg 200

Qy       1180 TGGCAGTCGATAAGTGGGCCCTGTATGTTAACGCACATATTCTTCCGAGTAACAGATTCG 1239
                                    :::   |||   ::::::   :::|||   :::
Db        201 -----------------------LeuAlaThrSerThrTyrTyrLeuLeuThrThrThr 212

Qy       1240 ---------------------TTCACAAAGGAGACCGTCGACAGTTTGTACAAATACCAC 1278
                                   ||||||   :::   |||   |||
Db        213 SerTyrLeuGlyMetLysCysAlaAsnLysGluAspPheGluTrpLeuThrLysAsnPro 232

Qy       1279 GATTTAGTTCGTTGGTCATCCTTCGTTCTGCGGCTTGCTGATGATTTGGGAACCTCGGTG 1338
                 ::::::   :::      :::   |||:::   ||||||:::   |||
Db        233 LysIleLeuGluAlaAsnValThrLeuCysArgValIleAspAspIleAlaThrTyrGlu 252

Qy       1339 GAAGAGGTGAGCAGAGGGGATGTGCCGAAATCACTTCAGTGCTACATGAGTGACTACAAT 1398
                 |||      ||||||   :::         ::::::|||||||||   ||||||
Db        253 ValGluLysGlyArgGlyGlnIleAlaThrGlyIleGluCysTyrMetArgAspTyrGly 272

Qy       1399 GCATCGGAGGCGGAGGCGCGGAAGCACGTGAAATGGCTGATAGCGGAGGTGTGGAAGAAG 1458
                 |||      ||||||   :::      :::   :::         ||||||
Db        273 ValSerThrGluGluAlaMetGluLysPheGlnGluMetAlaGluIleAlaTrpLysAsp 292

Qy       1459 ATGAATGCGGAGAGGGTGTCGAAGGATTCTCCATTCGGCAAAGATTTTATAGGATGTGCA 1518
              :::|||   |||   :::   :::   :::|||               :::
Db        293 ValAsn---GluGlyIleLeuArgProThrProValSerAlaLysIleLeuThrArgIle 311

Qy       1519 GTTGATTTAGGAAGGATGGCGCAGTTGATGTAC---CATAATGGAGATGGCACGGCACA 1575
              ::::::|||   |||:::      :::   |||   ||||||   ||||||:::
Db        312 LeuAsnLeuAlaArgIleIleAspValThrTyrLeuHisAsnGlnAspGlyTyr------ 329

Qy       1576 CAACACCCT 1584
                 ||||||
Db        330 ThrHisPro 332
```

Search completed: February 25, 2004, 05:50:38

Job time : 74 secs